

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY  
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

**ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ**

DEPARTMENT OF BIOMEDICAL ENGINEERING

**VYUŽITÍ STROJOVÉHO UČENÍ PRO GENEROVÁNÍ  
MEDICÍNSKÝCH OBRAZŮ**

MACHINE LEARNING BASED METHOD FOR MEDICAL IMAGE GENERATION

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**Valentina Hrtoňová**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. Roman Jakubíček, Ph.D.**

**BRNO 2021**

# Bakalářská práce

bakalářský studijní program **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

**Studentka:** Valentina Hrtoňová

**ID:** 203142

**Ročník:** 3

**Akademický rok:** 2020/21

**NÁZEV TÉMATU:**

## Využití strojového učení pro generování medicínských obrazů

### POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s problematikou generování syntetických obrazů s využitím strojového učení. Zejména se zaměřte na generativní soutěživé sítě (GAN). 2) Vypracujte rešerši na publikované metody a možnosti využití v medicíně. 3) Navrhněte obecné využití algoritmu, vyhledejte a připravte vhodné databáze libovolných 2D obrazů pro obecné testování. 4) Implementujte algoritmus ve vhodném programovém prostředí na zvolené databázi dat. 5) Navrhněte dle vypracované rešerše publikovaných metod generování syntetických obrazů pro medicínské účely. K tomu účelu vyhledejte a připravte vhodné databáze 2D obrazů. 6) Otestujte algoritmus na medicínských datech, diskutujte dosažené výsledky a možnosti využití.

### DOPORUČENÁ LITERATURA:

[1] GOODFELLOW, Ian J., Jean POUGET-ABADIE, Mehdi MIRZA, Bing XU, David WARDE-FARLEY, Sherjil OZAIR, Aaron COURVILLE a Yoshua BENGIO. Generative adversarial nets. In: Z. GHAHRAMANI, M. WELLING, C. CORTES, N. D. LAWRENCE a K. Q. WEINBERGER, ed. Advances in Neural Information Processing Systems [online]. B. m.: Curran Associates, Inc., 2014, s. 2672–2680. ISSN 10495258.

[2] YI, Xin, Ekta WALIA a Paul BABYN. Generative adversarial network in medical imaging: A review. Medical Image Analysis [online]. 2019, 58, 101552. ISSN 13618423.

**Termín zadání:** 8.2.2021

**Termín odevzdání:** 28.5.2021

**Vedoucí práce:** Ing. Roman Jakubíček, Ph.D.

**doc. Ing. Jana Kolářová, Ph.D.**  
předseda rady studijního programu

### UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Využití strojového učení pro generování medicínských obrazů“ jsem vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autorka uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....

podpis autorky

## PODĚKOVÁNÍ

Ráda bych poděkovala vedoucímu bakalářské práce panu Ing. Romanu Jakubíčkoví, Ph.D. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

# Obsah

<b>Úvod</b>	<b>1</b>
<b>1 Umělé neuronové sítě</b>	<b>2</b>
1.1 Biologický neuron . . . . .	2
1.2 Výpočetní model neuronu . . . . .	2
1.3 Vícevrstvé neuronové sítě . . . . .	4
1.4 Učení neuronové sítě . . . . .	5
1.5 Konvoluční neuronové sítě . . . . .	7
1.6 Architektura CNN . . . . .	8
<b>2 Generativní soutěživé sítě</b>	<b>13</b>
2.1 Trénování GAN . . . . .	14
2.2 Modifikace GAN . . . . .	15
2.3 Využití GAN v medicíně . . . . .	16
<b>3 Generování medicínských obrazů</b>	<b>19</b>
3.1 Databáze pro generativní soutěživé sítě . . . . .	19
3.2 Publikované medicínské aplikace . . . . .	19
<b>4 Experimentální data</b>	<b>22</b>
4.1 Dermatoskopické snímky kožních lézí . . . . .	22
4.2 Snímky CT páteře . . . . .	24
<b>5 Navržené modely pro generování medicínských obrazů</b>	<b>26</b>
5.1 Hluboce konvoluční GAN . . . . .	26
5.2 Model „pix2pix“ . . . . .	29
<b>6 Implementace algoritmů</b>	<b>32</b>
6.1 Hluboce konvoluční GAN . . . . .	32
6.2 Model „pix2pix“ . . . . .	34
<b>7 Diskuze dosažených výsledků</b>	<b>39</b>
7.1 Hluboce konvoluční GAN . . . . .	39
7.2 Model „pix2pix“ pro generování obrazů kožních lézí . . . . .	41
7.3 Model „pix2pix“ pro generování obrazů páteře . . . . .	45
7.4 Diskuze na téma generování medicínských obrazů modely GAN . . . . .	49
<b>Závěr</b>	<b>52</b>

<b>Literatura</b>	<b>53</b>
<b>Seznam zkratk</b>	<b>65</b>
<b>A Obsah elektronických příloh</b>	<b>66</b>

# Seznam obrázků

1.1	Neuron . . . . .	2
1.2	Perceptron . . . . .	3
1.3	Aktivační funkce . . . . .	4
1.4	Neuronová síť . . . . .	5
1.5	Konvoluční neuronová síť . . . . .	8
1.6	Konvoluční vrstva . . . . .	9
1.7	Aktivační funkce ReLU a lReLU . . . . .	10
1.8	Max a average pooling . . . . .	11
1.9	Dropout . . . . .	12
2.1	Generativní soutěživá síť . . . . .	13
3.1	Medicínská aplikace Baur et al. . . . .	20
3.2	Medicínská aplikace Galbusera et al. . . . .	21
3.3	Medicínská aplikace Nie et al. . . . .	21
4.1	ISIC 2020 Challenge Dataset . . . . .	22
4.2	ISIC 2017 Challenge Dataset . . . . .	24
4.3	Dataset CT páteře . . . . .	25
5.1	Generátor DCGAN . . . . .	27
5.2	Diskriminátor DCGAN . . . . .	28
5.3	Generátor typu „U-Net“ . . . . .	29
5.4	Diskriminátor typu „PatchGAN“ . . . . .	30
6.1	Blok generátoru – DCGAN . . . . .	33
6.2	Blok diskriminátoru – DCGAN . . . . .	33
6.3	Bloky generátoru – „pix2pix“ . . . . .	35
6.4	Blok diskriminátoru – „pix2pix“ . . . . .	36
6.5	Trénink „pix2pix“ – kožní léze . . . . .	37
6.6	Trénink „pix2pix“ – axiální snímky . . . . .	38
6.7	Trénink „pix2pix“ – sagitální snímky . . . . .	38
7.1	Kožní léze DCGAN – realistické . . . . .	39
7.2	Kožní léze DCGAN – nerealistické . . . . .	40
7.3	Kožní léze DCGAN – srovnání . . . . .	41
7.4	Kožní léze „pix2pix“ . . . . .	42
7.5	Kožní léze DCGAN – zajímavosti . . . . .	42
7.6	Kožní léze „pix2pix“ – srovnání . . . . .	43
7.7	Kožní léze – experiment . . . . .	44
7.8	Axiální řez . . . . .	46
7.9	Syntetický bederní obratel . . . . .	46
7.10	Axiální řezy – fyziologické masky . . . . .	47



7.11	Axiální řezy – patologické masky . . . . .	47
7.12	Axiální řezy – chybné . . . . .	48
7.13	Sagitální řez . . . . .	49

# Úvod

Generativní soutěživé sítě jsou speciálním typem umělých neuronových sítí, které v posledních letech dosahují nečekaně skvělých výsledků při generování syntetických obrazů. V medicínské oblasti se setkáváme s chronickým nedostatkem vyvážených a anotovaných souborů dat obsahujících dostatečné množství diagnostických snímků, jež by umožňovaly trénování přesných umělých neuronových sítí. Přístup k datům je často omezen pro zachování soukromí pacientů a snímky pozitivních případů vzácných patologií jsou v klinických databázích zastoupeny jen v malém množství. Generativní soutěživé sítě poskytují možné řešení tohoto problému a zároveň nacházejí uplatnění i v jiných oblastech analýzy medicínských obrazů jako jsou rekonstrukce, klasifikace, segmentace a detekce.

Teoretická část práce se v úvodu věnuje problematice umělých neuronových sítí, konkrétně jejich architektuře, učení neuronových sítí a podrobnějšímu rozboru konvolučních neuronových sítí. Je představen model generativních soutěživých sítí a je popsána jeho základní architektura, způsob trénování a nejvýznamnější modifikace. Dále je zpracována rešerše na možnosti využití generativních soutěživých sítí v oblasti medicíny a jsou prezentovány vybrané medicínské publikace na toto téma.

Praktická část práce je věnována návrhu, implementaci a testování několika modelů generativních soutěživých sítí. Prvním modelem je hluboce konvoluční generativní soutěživá síť aplikovaná na generování syntetických obrazů kožních lézí. Implementací tohoto modelu je zkoumána schopnost generativních soutěživých sítí syntetizovat medicínské obrazy pouze na základě vstupního vektoru šumu. Pro generování medicínských obrazů s vyšší kvalitou a klinickým významem je navržena modifikace generativní soutěživé sítě označovaná jako „pix2pix“. Tento model umožňuje generovat obrazy na základě informací zadaných formou vstupního obrazu a v praktické části je implementován pro tři různé aplikace. Cílem první aplikace je generování realistických obrazů kožních lézí na základě binární masky a porovnání výsledků s modelem hluboce konvoluční generativní soutěživé sítě. Pro další zkoumání podmíněné syntézy pomocí generativních soutěživých sítí je model „pix2pix“ aplikován s cílem generování syntetických CT obrazů páteře v axiální a sagitální rovině na základě binárních masek páteře.

V práci jsou podrobně popsány využití sady medicínských dat, návrh jednotlivých modelů a jejich implementace v programovém prostředí Python. Výsledky syntézy medicínských dat pomocí vybraných modelů jsou na závěr vyhodnoceny a diskutovány.

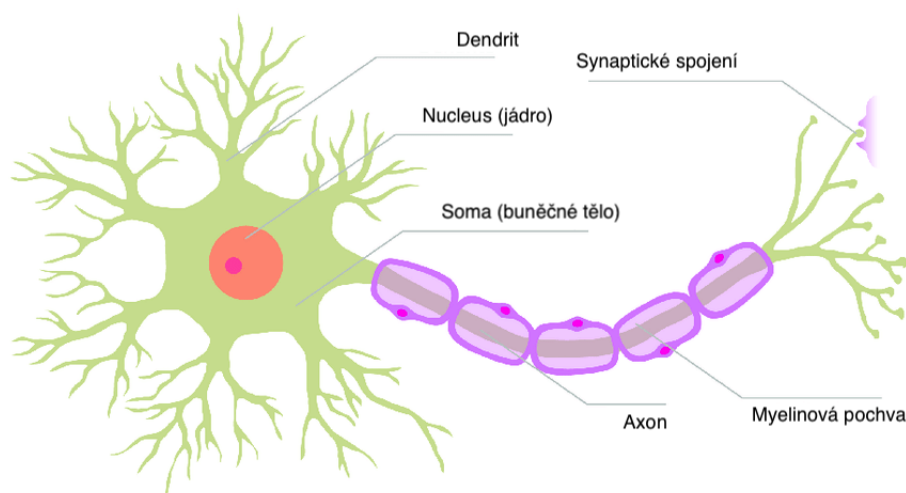
# 1 Umělé neuronové sítě

Umělé neuronové sítě jsou výpočetní sítě simulující rozhodovací proces, který pozorujeme v sítích nervových buněk (neuronů) biologického centrálního nervového systému.

## 1.1 Biologický neuron

Základní stavební a funkční jednotkou biologické nervové sítě zpracovávající informace je neuron, jehož struktura je ilustrována na obrázku 1.1.

Jádro neuronu je uloženo v buněčném těle (soma), k němuž jsou asociovány aferentní vlákna (dendrity), přijímající signály od sousedních neuronů a jedno dlouhé eferentní vlákno (axon), šířící signál směrem od těla k synapsím dalších neuronů. K přenosu informací ve formě akčního potenciálu dochází při překročení prahového napětí. [1]



Obr. 1.1: Ilustrace struktury fyziologického neuronu. Upraveno z [2].

## 1.2 Výpočetní model neuronu

Autoři McCulloch a Pitts navrhli v roce 1943 binární prahovou jednotku jako zjednodušený výpočetní model pro umělý neuron. Tento matematický neuron počítá vážený součet svých vstupních aktivačních nebo inhibičních signálů a generuje výstup 1, pokud tento součet překročí určitou prahovou hodnotou. Pokud se součet nachází pod prahovou hodnotou, výsledným výstupem je 0. [3]

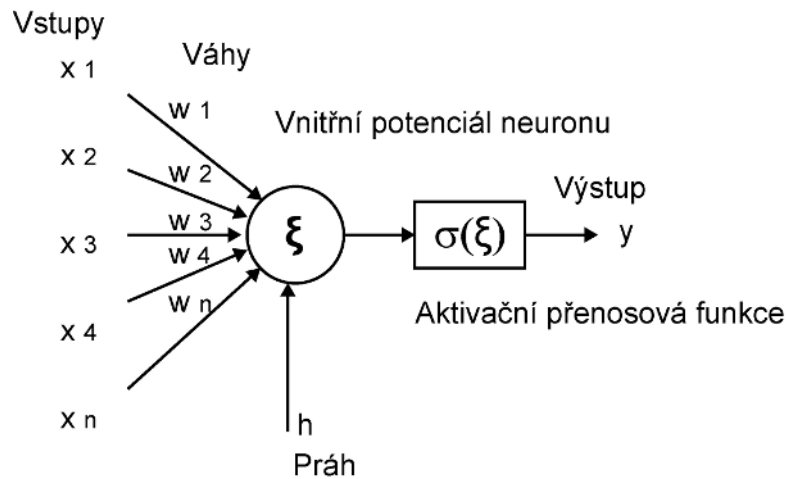
McCulloch-Pittsův model byl zobecněn a na jeho základě vyvinul na konci 50. let Frank Rosenblatt první algoritmus pro učení jednoduchých sítí, skládajících se z podobných modelů neuronu, tzv. perceptronů. Perceptron zachovává základní strukturu nervové buňky znázorněnou na 1.1 a lze ho rozdělit na několik částí [4]:

- synapse zatížené vahami  $w$ , modelující synaptické spojení, přijímající vstupní signály  $x$  a vedoucí je do těla neuronu,
- tělo neuronu, v němž je sumací získáván vnitřní potenciál neuronu  $\xi$ ,
- blok aktivační přenosové funkce  $\varphi$ ,
- výstupy neuronu  $y$  navazující na další perceptrony.

Do perceptronu vstupuje vektor vstupních hodnot  $x$  o  $n$  prvcích, vektor vah  $w$  také o  $n$  prvcích a práh  $h$ , často modelovaný jako jeden ze vstupů. Každý prvek vektoru hodnot je vynásoben příslušnou váhou, která vyjadřuje významnost jednotlivých vstupů. V těle neuronu proběhne sumace hodnot pro všech  $n$  prvků vstupního vektoru a od výsledné hodnoty je odečten práh  $h$ . Takto získaný vnitřní potenciál neuronu je nelineární aktivační funkcí transformován na výstupní hodnotu  $y$ .

Matematické chování neuronu je popsáno rovnicí 1.1 a model umělého neuronu znázorněn na obrázku 1.2.

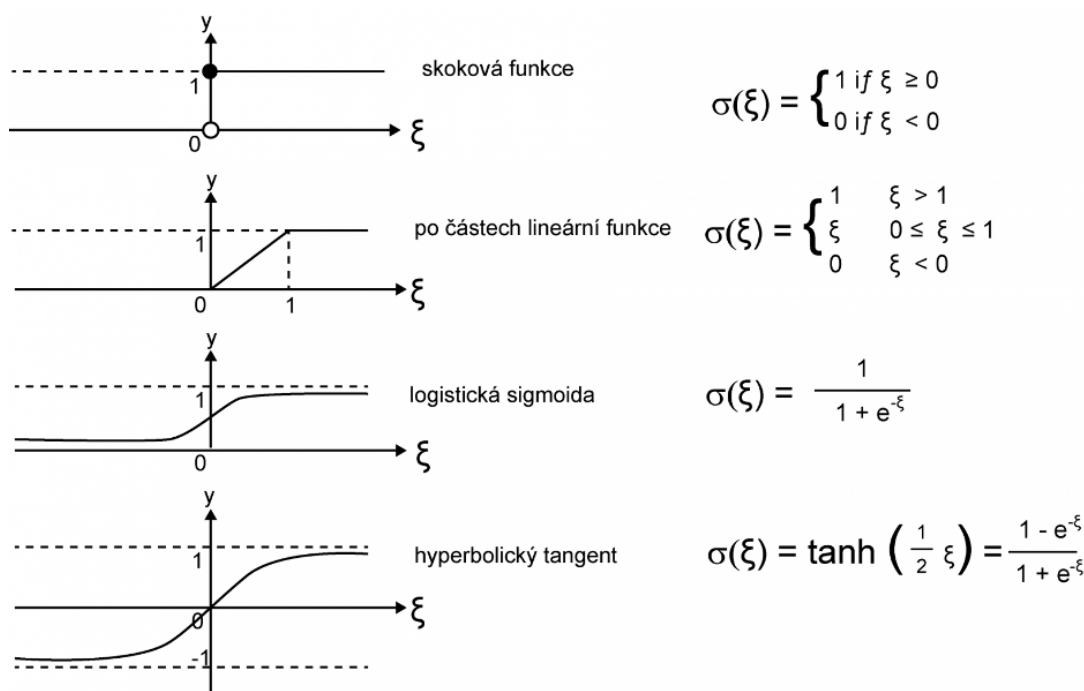
$$y = f\left[\sum_{i=1}^N w_i x_i - h\right] = f\left[\sum_{i=0}^N w_i x_i\right]; \quad x_0 = 1 \text{ a } w_0 = -h \quad (1.1)$$



Obr. 1.2: Schématické zobrazení umělého neuronu s  $n$  vstupy  $x$  (jimž přísluší váhy  $w$ ), prahem citlivosti neuronu  $h$ , aktivační přenosovou funkcí  $\varphi$  a jedním výstupem  $y$ . Převzato z [5].

Aktivační přenosová funkce je obecně nelineární funkcí transformující hodnotu vnitřního potenciálu neuronu. Základní aktivační funkce používané v dopředných

neuronových sítích jsou znázorněny a popsány na obrázku 1.3. Výběr funkce záleží na konkrétní aplikaci a požadavcích modelu.



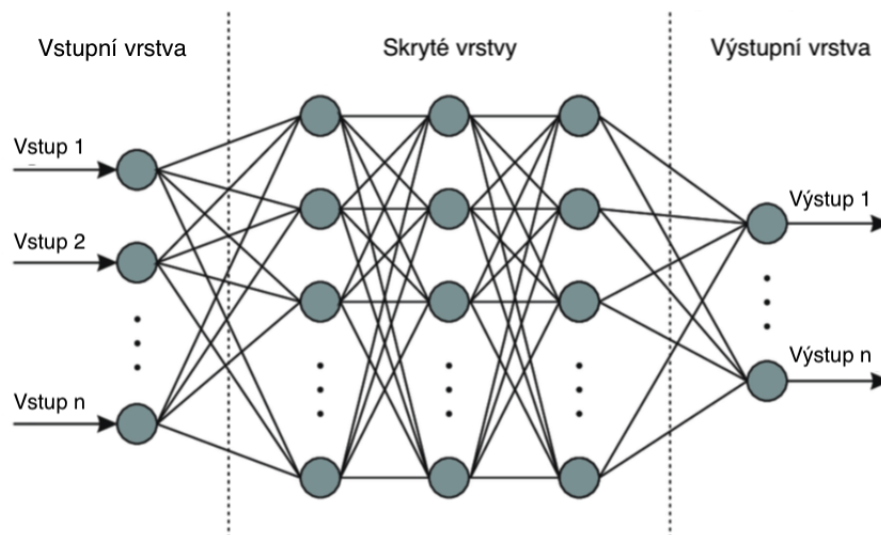
Obr. 1.3: Přehled základních aktivačních funkcí. Převzato z [5].

Perceptron nebo jednovrstvý perceptron lze použít na řešení jednoduchých problémů. Pro řešení složitějších úloh využíváme vícevrstvé perceptrony označované jako neuronové sítě.

## 1.3 Vícevrstvé neuronové sítě

Umělé neuronové sítě můžeme považovat za orientované grafy, v nichž mají neurony roli uzlů a orientované hrany hodnocené vahami jsou spojením mezi neuronovými výstupy a neuronovými vstupy reprezentujícím tok signálu [6]. Výkonné prvky vícevrstvé neuronové sítě (neurony) jsou uspořádány do minimálně tří vrstev: vstupní vrstvy, minimálně jedné skryté vrstvy a výstupní vrstvy. Příklad architektury vícevrstvé neuronové sítě je na obrázku 1.4.

Každý neuron vstupní vrstvy přijímá právě jeden vstupní signál, nepodílí se na výpočtu a pouze ho přivádí k dalším prvkům sítě. Neurony skrytých vrstev přenášejí a transformují signál přijatý vstupní vrstvou. Výstup každého z neuronů v dané vrstvě je přiváděn na vstupy všech neuronů ve vrstvě následující, přičemž neurony ve stejné vrstvě nejsou navzájem propojeny. Výsledkem výpočtu sítě, naposled zpracovaným výstupní vrstvou, je vektor výstupních hodnot výstupních neuronů. [8]



Obr. 1.4: Příklad architektury vícevrstvé neuronové sítě se vstupní vrstvou, třemi skrytými plně propojenými vrstvami a výstupní vrstvou. Upraveno z [7].

Spojení mezi neurony má vždy svůj směr a podle architektury dělíme neuronové sítě na dvě kategorie. Dopředné sítě neobsahují smyčky a obecně jsou statické, tj. produkují pouze jednu sadu výstupních hodnot místo sekvence hodnot z daného vstupu. Jejich odpověď na vstup je nezávislá na předchozích stavech sítě. Oproti tomu zpětnovazebné sítě obsahující smyčky jsou dynamickým systémem, v němž je výstup neuronu přepočítán s každým novým výstupem. Díky zpětné vazbě je s každým výstupem síť modifikována, což vede k novému stavu. Pro různé typy архитектур je třeba zvolit vhodné učící algoritmy. [3]

## 1.4 Učení neuronové sítě

Cílem učení neuronové sítě je nalézt takovou konfiguraci (hodnoty vah), aby síť co nejpřesněji realizovala transformaci zadanou tréninkovou množinou. Trénování neuronu probíhá iterativně ve várkách vzorků označovaných jako „mini-batch“ a interval, během kterého dojde k předložení všech vzorků z trénovací množiny, nazýváme epochou učení. K naučení sítě může být dle komplexnosti problému zapotřebí desítky až tisíce epoch. Učení může probíhat dvěma způsoby, a to s učitelem nebo bez učitele.

Při učení s učitelem je tréninková množina uspořádána na vektor vstupních hodnot sítě a na vektor požadovaných výstupních hodnot a je využita zpětná vazba. Během učení jsou odpovídající hodnoty z těchto dvou vektorů porovnávány a synaptické váhy sítě jsou upraveny na základě kritériální funkce. Při učení bez učitele

nemá optimalizační algoritmus k dispozici vektor na ověření správnosti transformace vstupních dat. Na základě poskytnutých vstupních dat a jejich společných elementů provádí shlukování do tříd. Učení bez učitele je vhodné využít pro dynamické sítě, jež na základě nového vzorku upravují současně i parametry sítě. Kombinací učení s učitelem a bez učitele je učení s částečným dohledem, při němž je k trénování modelu využita kombinace malého množství označených a velkého množství neoznačených dat.

Na počátku učicího procesu jsou inicializovány váhy (často náhodně) a pro výpočet jejich optimálních hodnot je využit optimalizační algoritmus. Pojem optimalizace vyjadřuje cíl minimalizovat nebo maximalizovat zvolenou kritériální funkci, jíž v případě minimalizace můžeme označit také jako chybovou funkci sítě [9]. Volba kritériální funkce musí odpovídat řešenému problému a konfiguraci výstupní vrstvy. Při učení s učitelem může být použita například střední kvadratická odchylka, kterou lze vypočítat podle vzorce:

$$L(y, \hat{y}) = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2, \quad (1.2)$$

kde  $\hat{y}_i$  je předpokládaný výstup sítě,  $y_i$  je skutečný výstup a  $m$  je počet vstupů.

Pro optimalizaci neuronových sítí jsou nejčastěji využívány algoritmy založené na gradientu.

### 1.4.1 Gradientní sestup

Základní gradientní metodou je metoda gradientního sestupu. Gradient, jenž je hodnotou derivace funkce v určitém bodě, určuje směr nejstrmějšího růstu kritériální funkce a pohybem proti gradientu malými kroky (definovanými jako učicí krok  $\epsilon$ ) jsme schopni kritériální funkci minimalizovat [9]. Tímto způsobem jsou upraveny váhy sítě dle vzorce:

$$w_{i+1} = w_i - \epsilon \frac{\partial L}{\partial w_i}, \quad (1.3)$$

kde  $w_{i+1}$  jsou aktualizované váhy,  $w_i$  jsou původní váhy,  $\epsilon$  je učicí krok a  $L$  je gradient kritériální funkce.

Pro rekurzivní výpočet gradientů skrytých vrstev vybraným optimalizačním algoritmem je využito zpětné šíření chyby (angl. „backpropagation“). Zpětné šíření chyby umožňuje aktualizovat všechny váhy současně během jednoho dopředného a zpětného průchodu sítí místo nutnosti opakovat proces pro každý jeden parametr. K úpravě vah dochází směrem od výstupní vrstvy k vrstvě vstupní, a tedy obráceným směrem, než kterým probíhá výpočet sítě v aktivním režimu. [8, 10]

Kriteriální funkce konverguje k lokálnímu nebo globálnímu minimu a proces se opakuje, dokud není dosažena jedna z podmínek ukončení učení. Základní metoda gradientního sestupu je však u nekonvexních funkcí náchylná k nežádoucímu uvíznutí v lokálním minimu a při učení sítě na velké datové sadě je také výpočetně náročná [9]. Proto jsou v praxi využívány její modifikace, jako je metoda stochastického gradientního sestupu.

### 1.4.2 Stochastický gradientní sestup

Metoda stochastického gradientního sestupu využívá pro optimalizaci vah průměrný gradient *mini-batch* namísto gradientu každého vzorku. U této metody je pro zajištění konvergence nutné postupně snižovat rychlost učení definovanou učícím krokem  $\epsilon$ . Důležitou vlastností stochastického gradientního sestupu je skutečnost, že doba potřebná pro aktualizaci parametrů v jedné iteraci neroste v závislosti na počtu vzorků v *mini-batch*, což umožňuje konvergenci i pro velmi velkou datovou sadu. U obsáhlých datových sad převažují rychlé počáteční výpočty gradientu tohoto algoritmu jeho pomalou asymptotickou konvergencí, čímž algoritmus umožňuje zrychlení učícího procesu. [9]

Tato metoda má mnoho dalších modifikací, řada z nichž se zaměřuje na adaptivní nastavení učícího kroku, který významně ovlivňuje výkon modelu.

### 1.4.3 Metoda adaptivního odhadu momentu

Jedním z rozšíření stochastické metody gradientního sestupu je metoda adaptivního odhadu momentu (Adam). Hlavní výhodou metody Adam je schopnost adaptivních změn učícího kroku v průběhu trénování sítě. Úprava parametrů sítě pomocí metody Adam je definována učícím krokem  $\alpha$  a mírou exponenciálního zapomínání pro odhady momentů  $\beta_1, \beta_2$ . [11]

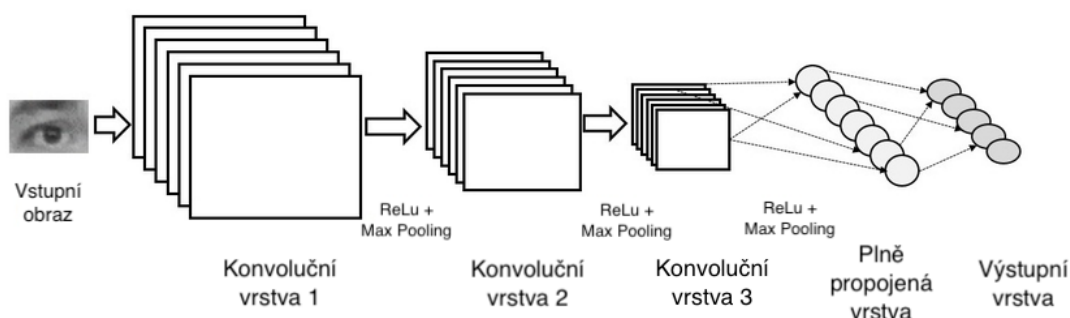
## 1.5 Konvoluční neuronové sítě

Vysoce efektivním modelem pro zpracování obrazových dat jsou konvoluční neuronové sítě (CNN), jež dosahují výborných výsledků v rozpoznávání vzorů například při klasifikaci obsáhlých datových sad, počítačové vidění nebo také při počítačové zpracování přirozeného jazyka. CNN jsou schopny se automaticky a adaptivně učit hierarchie příznaků, a to pomocí metody zpětného šíření chyby a za využití několika základních stavebních bloků, jako jsou konvoluční vrstvy, podvzorkovací vrstvy a plně propojené vrstvy.



## 1.6 Architektura CNN

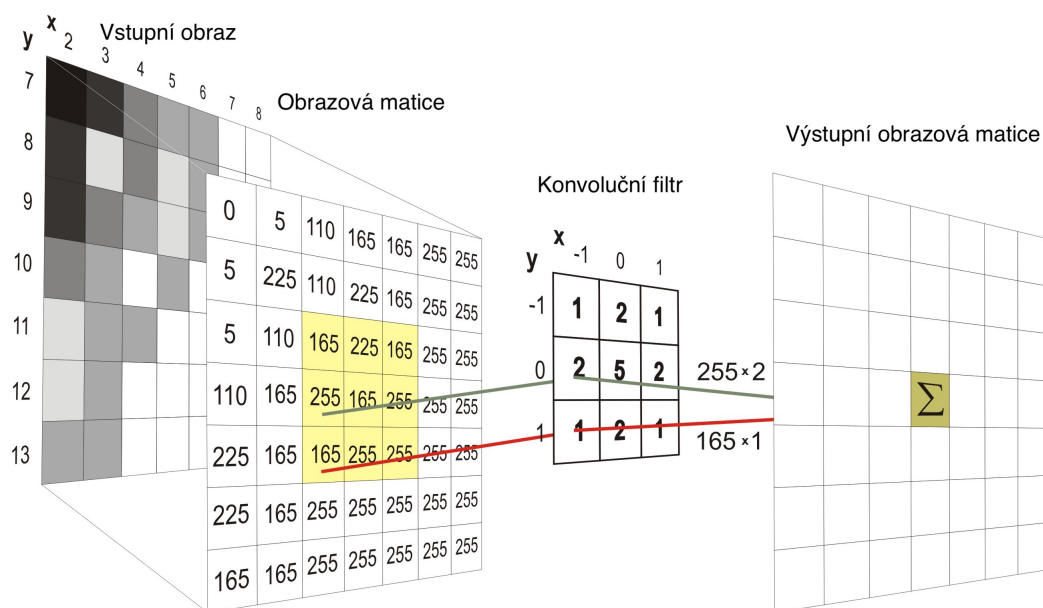
CNN lze rozdělit na síť provádějící extrakci příznaků a na klasifikační nebo regresní síť, která tyto příznaky vyhodnocuje a generuje výstup. Příklad architektury CNN je na obrázku 1.5. První část sítě je tvořena opakujícími se bloky konvolučních a podvzorkovacích vrstev a zodpovídá za extrakci příznaků z obrazových dat. Se stoupajícím počtem bloků jsou zachyceny detailnější příznaky a vzniklé příznakové mapy jsou v druhé části sítě zpracovávány plně propojenými vrstvami, které provádí klasifikaci nebo regresi. V následujících podkapitolách jsou popsány základní vrstvy CNN a další vrstvy CNN využívané v architektuře generativních soutěživých sítí.



Obr. 1.5: Příklad architektury sítě typu CNN. Vstupní vrstva je následována třemi bloky obsahujícími: konvoluční vrstvu, aktivační funkci typu „Rectifier Linear Unit“ (ReLU) a podvzorkovací vrstvu „max pooling“. Na poslední blok navazuje plně propojená vrstva a výstupní vrstva. Upraveno z [12].

### 1.6.1 Konvoluční vrstva

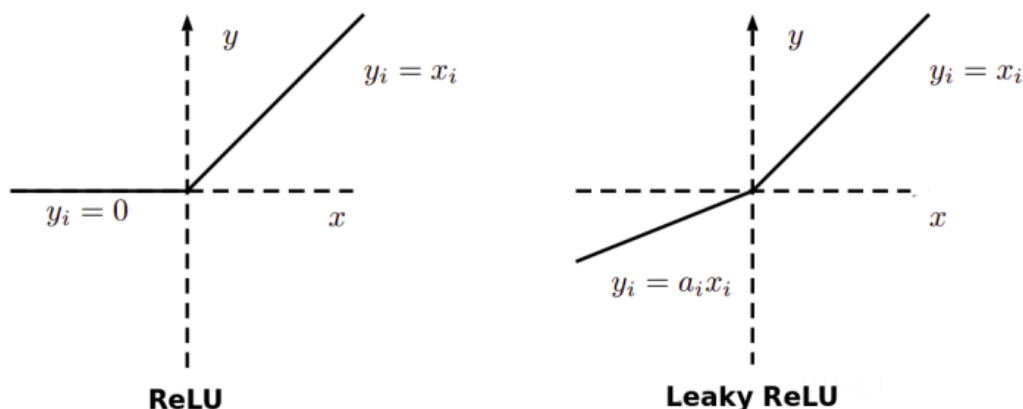
Konvoluční vrstvy, které jsou typicky tvořeny kombinací lineárních operací, získávají pomocí konvolučních filtrů lokální informace o vstupním obrazu ve formátu příznakových map. Konvoluční filtry (nebo také konvoluční masky) jsou dvojrozměrné matice o relativně malém rozměru (velikosti jádra), většinou  $3 \times 3$  nebo  $5 \times 5$  a jejich hodnoty jsou optimalizovány během procesu učení. Tento proces je obdobný jako aktualizace vah u běžné umělé neuronové sítě. Při průchodu konvoluční vrstvou je provedena konvoluce daného filtru se vstupní maticí, jak je zobrazeno na obrázku 1.6. Tento proces se opakuje pro každý filtr a vytvářejí se příznakové mapy reprezentující různé charakteristiky vstupní matice, u obrazových dat například hrany, čáry a detailnější vizuální prvky. Různé konvoluční filtry mohou tedy být považovány za extraktory různých příznaků a jejich počet určuje počet výstupních příznakových map, a tím hloubku vrstvy [14].



Obr. 1.6: Princip výpočtu konvoluce konvolučním filtrem s velikostí jádra (3,3). Výstupní obrazová matice je vyjádřením příznakové mapy. Upraveno z [13].

Během validní konvoluce dochází ke zmenšení rozměrů vstupní matice. Pro řešení tohoto problému se používá metoda označovaná jako „zero padding“, při níž jsou na každou stranu vstupu přidány řady a sloupce nul a díky tomu je zachován původní rozměr matice. Dalším způsobem, jak definovat konvoluční operaci, je nastavení hyperparametru **stride**, který vyjadřuje vzdálenost mezi dvěma po sobě následujícími pozicemi recepčního pole. Nastavením **stride** na hodnotu větší než 1 lze snížit překryv a docílit podvzorkování příznakové mapy, které je jinak běžně prováděno podvzorkovací vrstvou. Další klíčovou vlastností konvolučních vrstev je sdílení vah. Konvoluční filtry, jež jsou přikládány na vstup, jsou charakterizované těmito vahami a jsou neměnné na všech pozicích vstupu. Díky sdílení vah je značně omezen počet parametrů sítě, což zvyšuje její efektivitu a lokální propojenost umožňuje síti detekovat a rozpoznávat prvky bez ohledu na jejich polohu v obraze. [14]

Konvoluční vrstva je většinou následována vrstvou zavádějící do systému nelinearit a modelující vygenerovaný výstup. V současnosti je pro tuto aplikaci nejpožívanější aktivační funkce „Rectifier Linear Unit“ (ReLU), která záporné hodnoty transformuje na hodnotu 0 a nezáporné ponechá beze změny. Jednou z modifikací této funkce je aktivační funkce „Leaky Rectifier Linear Unit“ (lReLU), jež nenuluje záporné vstupy, ale násobí je konstantním gradientem  $\alpha$ . Průběh obou funkcí je na obrázku 1.7.



Obr. 1.7: Aktivační funkce ReLU a lReLU. Převzato z [15].

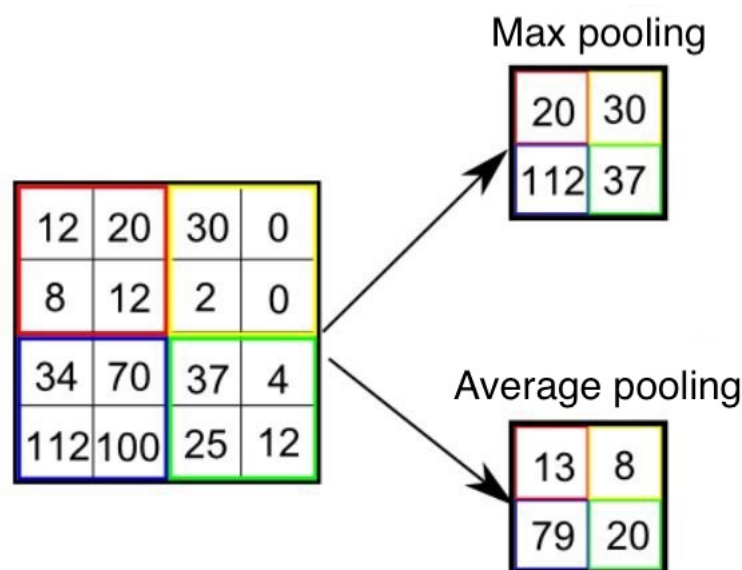
### 1.6.2 Podvzorkovací vrstva

Úlohou podvzorkovacích vrstev (z angl. „pooling“) v konvoluční síti je progresivní snižování complexity vrstev, a tím množství parametrů sítě. Vrstva provádí podvzorkovací operaci, během níž jsou sloučeny sousedící pixely určité oblasti do jedné reprezentativní hodnoty. Tato hodnota je většinou vypočítána jako maximum („max pooling“) nebo průměr („average pooling“) vybraných pixelů. [16]

V praxi je pro podvzorkování běžně používána maska o velikosti  $2 \times 2$  s hodnotou **stride** 2, což vede ke zmenšení prostorové dimenze (rozlišení) na polovinu [17]. Příklad podvzorkování metodami „max pooling“ a „average pooling“ je na obrázku 1.8.

### 1.6.3 Plně propojená vrstva

Příznakové mapy získané sítí provádějící extrakci příznaků jsou dále zpracovány druhou sítí s plně propojenými vrstvami. Výstup posledního bloku konvoluční a podvzorkovací vrstvy je transformován do vektoru hodnot a přiveden k jedné nebo více plně propojeným vrstvám. Každý neuron v plně propojené vrstvě je váhovaným spojením přímo spojen s každým neuronem v předchozí i následující vrstvě a konečným výstupem této sítě pro klasifikační úlohy jsou pravděpodobnosti příslušnosti do jednotlivých tříd. Poslední plně propojená vrstva má obvykle stejný počet neuronů jako je počet tříd a v závislosti na řešené úloze je vybrána vhodná aktivační funkce. Pro klasifikační úlohu do více tříd je vhodnou aktivační funkcí např. funkce „softmax“, jež normalizuje výstupy tak, aby jejich celkový součet byl 1. [14]



Obr. 1.8: Ukázka podvzorkovacích metod „max pooling“ a „average pooling“ za použití masky  $2 \times 2$  a nastavení hyperparametru  $\text{stride} = 2$ . Upraveno z [18].

#### 1.6.4 Vrstva transponované konvoluce

Některé aplikace CNN, jako je sémantická segmentace nebo generativní soutěživé sítě, vyžadují nadvzorkování obrazové matice, během něž je zvětšena její prostorová dimenze. Nejčastěji je pro tyto účely využita vrstva transponované konvoluce, jež provádí inverzní konvoluční operaci. Vrstva vstup smysluplně transformuje a při správném nastavení hyperparametru `stride` zvyšuje rozlišení obrazu. Nastavením `stride` na hodnotu 2 u konvoluční vrstvy dosáhneme podvzorkování vstupu a zmenšení prostorové dimenze na polovinu, naopak u vrstvy transponované konvoluce inverzně ke zvětšení této dimenze ve stejné míře. Parametry této vrstvy, jež jsou využity při rekonstrukci původních vstupů, jsou během tréninku optimalizovány společně s parametry konvolučních vrstev. [19]

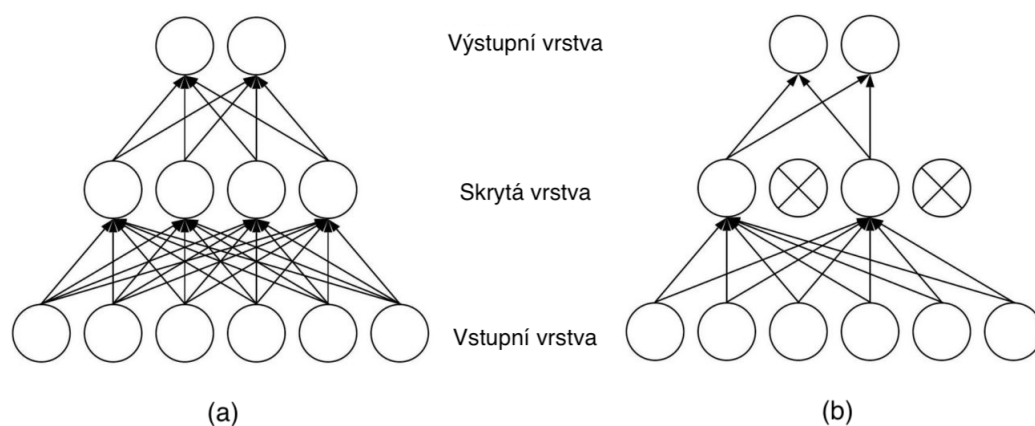
#### 1.6.5 Regularizační vrstvy

Jako opatření proti přeučení sítě (angl. „overfitting“) a při regularizaci tréninkového procesu je možné využít vrstvy „Batch Normalization“ a „Dropout“.

Vrstva *Batch Normalization* normalizuje hodnoty každé *mini-batch* tak, aby měly průměr 0 a směrodatnou odchylku 1. V praxi *Batch Normalization* nejčastěji normalizuje výstupy konvoluční vrstvy a je následována aktivační vrstvou. Působí jako zdroj šumu a zároveň umožňuje využití vyššího učícího kroku při trénování sítě. Během tréninkové fáze ukládá tato vrstva statistiky o svých vstupech, jež jsou poté

využity během testování. [20]

Vrstva *Dropout* implementuje stejnojmennou metodu založenou na dočasném vypnutí části neuronů, ilustrovanou na obrázku 1.9. Během tréninku jsou náhodně vybrané neurony dočasně odstraněny ze sítě s předem danou pravděpodobností. Tato pravděpodobnost je určena parametrem  $p$ , jež je běžně nastaven na hodnotu 0,5 odpovídající 50% pravděpodobnosti vypnutí neuronu. Pomocí této metody je předcházeno přeučení sítě a je snížen počet parametrů modelu. [21]



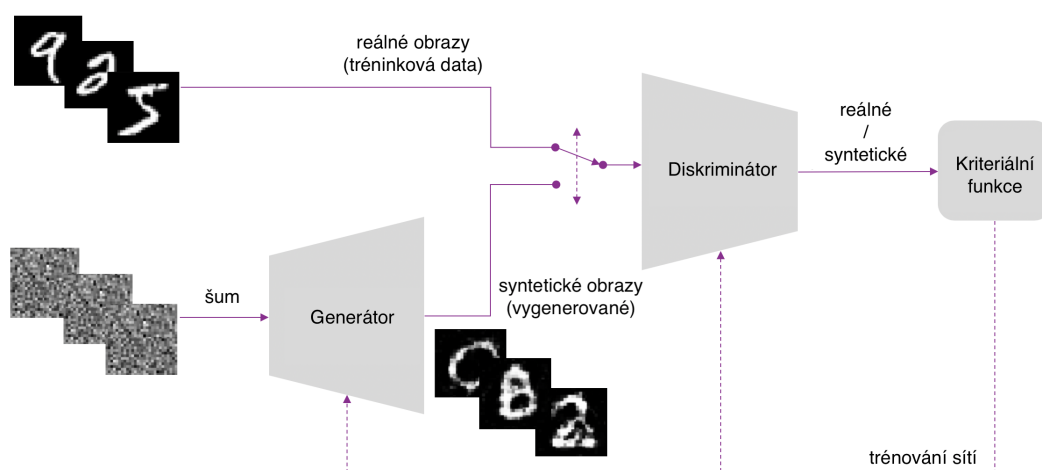
Obr. 1.9: Srovnání běžné sítě s plně propojenými vrstvami (a) a sítě, u níž je na skrytou vrstvu aplikován *Dropout* (b). Upraveno z [22].

## 2 Generativní soutěživé sítě

Generativní soutěživé sítě (GAN) jsou speciálním typem neuronových sítí, v nichž jsou trénovány dvě sítě současně, přičemž jedna je zaměřena na generování obrazu a druhá na diskriminaci.

Architektura GAN byla poprvé popsána v roce 2014 ve článku Iana Goodfellowa et al. s názvem „Generative Adversarial Networks” [23] a zahrnuje dva soutěžící modely: model generátoru pro generování syntetických vzorků a model diskriminátoru, který klasifikuje vzorky jako reálné (z množiny tréninkových dat) nebo syntetické (vytvořené generátorem). Schéma základního modelu GAN je na obrázku 2.1.

Generativní model generuje z prostoru s Gaussovou distribucí náhodný vektor a využívá ho pro započetí generativního procesu. Tento vektorový prostor, označovaný také jako „latentní prostor“, je složen z latentních proměnných, jež nelze přímo pozorovat. Latentní prostor se trénováním modelu stává projekcí trénovacích dat a poskytuje informace o jejich distribuci, na základě nichž je model generátoru schopen tvořit různé výstupy s podobnými charakteristikami, jako mají tato data. Model GAN může být tvořen neuronovou sítí s libovolnou strukturou za předpokladu, že rozměry vstupu a výstupu odpovídají rozměrům latentního prostoru a tréninkových dat. Diskriminativní model je binárním klasifikátorem, jenž zpracovává vzorky z množiny tréninkových a vygenerovaných dat a předpovídá, do které třídy vzorek patří. [25, 26]



Obr. 2.1: Blokové schéma základního modelu GAN. Upraveno z [24].

## 2.1 Trénování GAN

Model generátoru  $G$  využívá náhodnou realizaci šumu  $z$  („latentní vektor“), pocházející z distribuce  $p(z)$  jako vstup. Tento *latentní vektor* je prostřednictvím  $G$  transformován na vzorek  $x_g$ , u nějž předpokládáme vizuální podobnost s reálným vzorkem  $x_r$  z distribuce reálných dat  $p_r(x)$ . Primárním cílem  $G$  je optimalizovat mapování z *latentního prostoru* tak, aby se distribuce generovaných dat  $p_g(x)$  co nejvíce podobala distribuci reálných dat  $p_r(x)$  a po úspěšném tréninku mohla být považována za její aproximaci. K dosažení tohoto cíle slouží model diskriminátoru  $D$ , který klasifikuje mezi reálnými a syntetickými daty. Výstupem  $D$  je pravděpodobnost, s níž vzorek pochází z reálné distribuce dat  $p_r(x)$ , spíše než z distribuce syntetických dat  $p_g(x)$ . [23, 27]

Učící proces generátoru a diskriminátoru lze matematicky vyjádřit rovnicí:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_r(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (2.1)$$

kde  $D(x)$  je výstup diskriminátoru,  $G(z)$  výstup generátoru,  $p_r(x)$  reálná distribuce dat a  $p_z(z)$  distribuce *latentního prostoru* [23]. Model diskriminátoru a model generátoru jsou střídavě trénovány, procesem označovaným jako „minimax hra“, během nějž se snažíme dosáhnout rovnováhy mezi oběma modely. Cílem  $D$  je maximalizovat pravděpodobnost, že správně klasifikuje vzorky z distribuce  $p_r(x)$  jako reálné a vzorky z distribuce  $p_g(x)$  jako syntetické. Naopak  $G$  se snaží minimalizovat pravděpodobnost, že diskriminátor klasifikuje jeho výstupy jako syntetické.

Kriteriální funkci diskriminátoru lze matematicky definovat vzorcem 2.2 a kriteriální funkci generátoru vzorcem 2.3 [27].

$$\mathcal{L}_D^{GAN} = -\mathbb{E}_{x \sim p_r(x)} [\log D(x)] - \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2.2)$$

$$\mathcal{L}_G^{GAN} = \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2.3)$$

Modely  $G$  a  $D$  jsou v praxi střídavě trénovány metodou stochastického gradientního sestupu. Proces tréninku každé ze sítí je jiný a trénink celého modelu GAN probíhá ve střídajících se periodách. Během jedné periody probíhá  $k$  kroků optimalizace  $D$  a jeden krok optimalizace  $G$ . Diskriminátor je penalizován za špatnou klasifikaci vzorků a generátor za vytvoření vzorku, který  $D$  klasifikoval jako syntetický. Požadovaným výsledkem trénování GAN je, aby vzorky  $x_g$  generované  $G$  aproximovaly reálnou distribuci dat  $p_r(x)$ . V takovéto fázi je možné  $G$  využít pro úspěšné generování nových vzorků. [23, 28]

Během trénování GAN ovšem narážíme na několik komplikací. Při střídavém trénování  $G$  a  $D$  nelze zaručit rovnováhu a jedna ze sítí, převážně  $D$ , se hned z počátku tréninku stává „silnější“ než síť druhá. Pokud dokáže  $D$  příliš snadno rozeznat generované vzorky od skutečných, dochází k tzv. saturaci kritériální funkce, kdy se gradienty  $D$  blíží nule a neumožňují další trénink  $G$ . Pro trénink generátoru byla proto navržena upravená forma kritériální funkce 2.3, jež se místo trénování generátoru s cílem minimalizace  $\log(1 - D(G(z)))$  snaží maximalizovat  $\log(D(G(z)))$ . Jinými slovy se generátor snaží maximalizovat pravděpodobnost, že diskriminátor špatně klasifikuje ním vytvořené vzorky. Tato úprava nijak nemění dynamiku diskriminátoru a generátoru, předchází však saturaci  $\log(1 - D(G(z)))$  tím, že z počátku tréninku poskytuje silnější gradienty pro úpravu parametrů generátoru. [23]

Dalším problémem, k němuž dochází při trénování modelu GAN, je tzv. kolaps režimu generátoru („mode collapse“), během nějž se distribuce  $p_g(x)$  vytvořená  $G$  zaměřuje pouze na několik omezených režimů distribuce dat  $p_r(x)$  [29]. Místo vytváření různorodých obrazů generuje  $G$  jen omezenou sadu vzorků a modely nekonvergují.

S cílem stabilizovat trénink a také zabránit kolapsu režimu generátoru byly navrženy různé modifikace původního modelu GAN. Některé z těchto modifikací jsou popsány v následující kapitole.

## 2.2 Modifikace GAN

V počátečním nastavení byly v GAN používány plně propojené vrstvy a generování dat nebylo omezeno. Později byla však tato struktura nahrazena plně konvolučními vrstvami a poskytnutí pomocných informací generátoru umožnilo získání obrazů s požadovanými vlastnostmi. Pro splnění požadovaného výstupu bylo navrženo mnoho různých variant modelu GAN, některé z nich jsou popsány níže. [30]

Hluboce konvoluční GAN (DCGAN) [31], v jehož architektuře jsou plně propojené vrstvy nahrazeny plně konvolučními vrstvami, vykazuje lepší výsledky a větší stabilitu během trénovacího procesu. Konvoluční vrstva je v tomto modelu následována vrstvou *Batch Normalization*, jež normalizuje výstup z konvoluční vrstvy, čímž zabraňuje zhroucení režimu generátoru a aktivační funkcí typu lReLU. Aktivační funkce tohoto typu zlepšuje kvalitu obrazového výstupu a umožňuje generovat obrazy s vyšším rozlišením. [29]

V původním modelu GAN není možné ovlivňovat samotné generování dat, proto byl navržen model podmíněného GAN (cGAN) [32], jenž umožňuje řízení generativního procesu pomocí dalších vstupů. Cílem generátoru je vygenerovat obrazy, které jsou realistické a zároveň odpovídají zadanému kontextu. Jako vstup, na jehož základě cGAN generuje a rozlišuje obrazy, lze použít například diskrétní značky, text nebo obraz. Pokud je namísto náhodného šumu jako vstup generátoru cGAN



použit obraz, označujeme síť jako „pix2pix“ [33]. Model *pix2pix* lze využít pro transformaci obrazu na obraz („image-to-image“ transformace) a tím generování obrazů s vysokým rozlišením. Dalším podmíněným modelem GAN je InfoGAN [34], který se učí bez učitele a místo použití jediného nestrukturovaného šumového vektoru je v InfoGAN vektor šumu rozdělen na dvě části: šum a latentní kód zaměřující se na významné rysy tréninkových dat.

Velmi efektivním se při generování obrazů s vysokým rozlišením prokázalo také využití progresivních modelů jako jsou „Laplacian GAN“ (LAPGAN) [35], „Stacked GAN“ (StackGAN) [36] a „Progressive Growing GAN“ (PGAN) [37]. Model LAPGAN obsahuje několik na sebe navazujících modelů cGAN, z nichž každý přidává do generovaného obrazu podrobnosti o vyšší frekvenci. Kaskádu sítí GAN využívá ve své architektuře také StackGAN, jenž je rozložením generativního procesu na menší části schopné z textu generovat obrazy s vysokým rozlišením. V modelu PGAN progresivně rostou sítě generátoru a diskriminátoru tím, že je v průběhu trénování zvyšován počet jejich vrstev, díky čemuž jsou v obraze generovány detailnější struktury.

## 2.3 Využití GAN v medicíně

V medicíně se v současnosti GAN uplatňují dvěma způsoby. První se zaměřuje na generativní aspekt GAN, který pomáhá zkoumat a objevovat základní strukturu tréninkových dat a učí se generovat nové obrazy. Druhý využívá diskriminativní část tohoto modelu, kdy je diskriminátor naučen rozpoznávat „normální“ snímky a může být využit jako detektor abnormalit a klasifikátor. Metody založené na GAN byly použity pro úpravu snímků z celé řady lékařských zobrazovacích metod, jako je magnetická rezonance (MRI), výpočetní tomografie (CT), optická koherenční tomografie, rentgen (RTG), dermatoskopie, ultrazvuk, pozitronová emisní tomografie (PET) nebo mikroskopie [28].

### 2.3.1 Syntéza

Modely GAN jsou široce využívány pro syntézu medicínských obrazů, čímž pomáhají překonat problémy s ochranou osobních údajů související s diagnostickými lékařskými obrazovými daty a nedostatek snímků pozitivních případů vzácných patologií. Poskytují v porovnání s tradičními metodami (změna měřítka, rotace, translace a deformace) sofistikovanější řešení a byly použity v mnoha publikacích pro rozšíření souborů tréninkových snímků se slibnými výsledky. Původní GAN modely byly navrženy pro nekontrolované generování syntetického obrazu z latentního prostoru, avšak s příchodem cGAN se úspěšně vyvinuly do generativního modelu učícího se

s učitelem, který dokáže generovat data podle zadaných informací. Syntézu obrazu pomocí GAN tak můžeme rozdělit na dvě kategorie: nepodmíněná syntéza obrazu a podmíněná syntéza obrazu.

Pomocí modelů DCGAN [38, 39], PGAN [40] a LAPGAN [41], jež dokážou nepodmíněně syntetizovat data, lze řešit problémy jako je nedostatek dat nebo nevyváženost anotovaných tříd dat. Takováto syntéza pomáhá také získat hlubší vhled do podstaty distribuce dat a jejich latentní struktury.

Naopak modely pro podmíněnou syntézu umožňují syntetizovat umělé lékařské obrazy, jež mohou být podmíněny dalším vstupním obrazem, segmentačními mapami nebo textem. Lze tak syntetizovat například uzliny dotýkající se hranice plic pro efektivnější automatickou segmentaci patologií na plicních CT snímcích [42], barevný obraz sítnice ze syntetického cévního stromu [43] nebo realistické RTG snímky s různými charakteristickými onemocněními [44]. Speciálním druhem podmíněné syntézy je pak intermodální syntéza poskytující uměle generované obrazy cílové modalit z dostupné modalit se zachováním anatomické struktury a vlastností bez nutnosti samostatného získávání obou lékařských snímků. Lze syntetizovat snímky MRI na základě vstupních CT snímků [45] a naopak CT snímky na základě snímků MRI, čehož může být využito při plánování radioterapie [47]. Je možné generovat také snímky PET na základě CT snímků [48] a MRI snímků [49].

### 2.3.2 Rekonstrukce

Kvůli překážkám v klinickém prostředí, jako je dávka záření a pohodlí pacienta, může být diagnostická kvalita získaných lékařských snímků omezena nízkým prostorovým rozlišením, výskytem artefaktů nebo kontaminací šumem. Množství šumu v CT snímcích je možné redukovat například pomocí cGAN, jež dokáže reprodukovat CT snímky pořízené s rutinní dávkou ionizační energie („routine-dose“) na základě snímků pořízených s nízkou dávkou („low-dose“) [50]. Další využití nacházejí modifikace GAN při rekonstrukci MRI snímků. Například „Wasserstein GAN“ (WGAN) byl úspěšně využit pro trénování víceúrovňové hustě propojené sítě se super rozlišením [51]. Navržená metoda dokáže obnovit místní textury a detaily MRI snímku s nízkým rozlišením na úroveň snímku s vysokým rozlišením. Modifikovaným GAN, nazvaným „3D-cGAN“, lze zlepšit kvalitu snímků PET, a tím snížit radiační zátěž při zachování vysoké kvality těchto snímků [52].

### 2.3.3 Klasifikace

Pro řešení klasifikace využívají lékařské zobrazovací studie obvykle konvoluční neuronové sítě učící se s učitelem, které ovšem k dosažení dobrých výsledků vyžadují velké datové sady anotované odborníky. Pro řešení klasifikačních problémů však

mohou být využity i GAN. Části generátoru a diskriminátoru mohou být využity jako extraktor příznaků nebo přímo diskriminativní síť jako klasifikátor, jenž každému obrazu přiřadí třídu [29]. Hu et al. [53] představili metodu učení bez učitele inspirovanou WGAN a InfoGAN, jež je schopna klasifikovat buněčné složky kostní dřeně podle jejich nejvýznamnějších rysů. Modely GAN byly dále využity například pro klasifikaci abnormalit ze snímků sítnicové fundoskopie [54], detekci a klasifikaci melanomu z dermatoskopických snímků [55] a pro klasifikaci srdečních abnormalit z RTG snímků hrudníku [56].

### 2.3.4 Další využití

Diskriminační část modelu GAN lze využít pro segmentaci a detekci abnormalit. Diskriminátor může být v tomto případě považována za regulátor tvaru a je schopen segmentovat dostatečně kompaktní objekty, jako jsou srdce, plíce nebo játra. Například pro automatickou segmentaci jater z CT a MRI snímků byl navržen přístup pracující se soutěživou „image-to-image“ sítí inspirovanou GAN [57]. Učením se distribuce tréninkových dat zobrazujících fyziologické snímky je diskriminátor schopen detekovat obrazy, které do této distribuce nepatří a které lze považovat za abnormální. Schlegl et al. [58] ve své studii prokázali úspěšné využití hluboce konvolučního GAN („AnoGAN“) pro detekci anomálií sítnice ze snímků optické koherentní tomografie a Styner et al. [59] použili model GAN pro detekci mozkových lézí na MRI snímcích.

Dále byly modely GAN použity pro modelování specifické distribuce pohybu pacienta na základě předoperačního obrazu [60], zbarvení endoskopických video dat [61] nebo plánování léčby v radioterapii předpovědí distribuce dávky z CT obrazu [62].

## 3 Generování medicínských obrazů

Výkonnost modelů pro rozpoznávání lékařských snímků velmi závisí na reprezentativnosti a dostatečné četnosti tréninkových vzorků. Vysoké náklady na sběr velkého množství praktických lékařských obrazů vedou k poptávce po syntéze obrazových vzorků např. za využití GAN. Model GAN byl navržen k překonání nevýhod jiných generativních algoritmů, jejichž příklady a srovnání s GAN lze nalézt v této publikaci [63].

### 3.1 Databáze pro generativní soutěživé sítě

V současnosti probíhají snahy o vytváření velkých otevřených databází jako např. Biobank, The National Biomedical Imaging Archive, The Cancer Imaging Archive and Radiologist Society of North America, ovšem nedostatek vhodných tréninkových dat zůstává nadále komplikací ve výzkumu a aplikaci GAN. [29]

### 3.2 Publikované medicínské aplikace

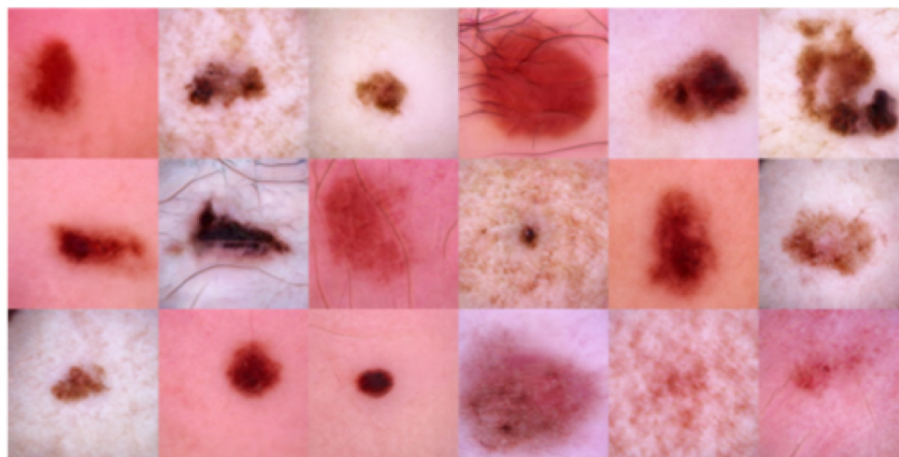
Generování medicínských obrazů pomocí GAN lze rozdělit do dvou hlavních kategorií: nepodmíněné a podmíněné syntézy. Kategorie podmíněné syntézy zahrnuje ještě speciální skupinu, a to intermodální syntézu. Z řady publikovaných vědeckých prací byly v práci popsány tři studie, každá spadající do jedné ze tří kategorií.

Práce týmu Baur et al. [64] na syntéze obrazů kožních lézí byla motivována nedostatkem anotovaných dat a zkrácením zastoupením sedmi kategorií lézí v dostupných datech. Takováto data neumožňují trénování robustních modelů schopných úspěšné klasifikace a segmentace lézí, a proto bylo jako řešení navrženo použití GAN, konkrétně jeho progresivní modifikace PGAN. Jako zdrojová data byly použity dermatoskopické snímky<sup>1</sup> benigních a maligních kožních lézí poskytnuté v rámci výzvy ISIC2018 [66], na jejichž základě byl PGAN schopen vygenerovat velmi realistické syntetické obrazy v rozlišení  $256 \times 256$  pixelů. Tyto obrazy byly odbornými dermatology během vizuálního Turingova testu jen obtížně odlišitelné od reálných. Ve studii byly dále srovnány obrazy syntetizované PGAN s výstupy modelů DCGAN a LAPGAN, které se osvědčily při syntéze vysoce kvalitního obrazu ze šumu, avšak při nízkém rozlišení. Pro kvantitativní posouzení reálnosti obrazů byla použita metrika „Sliced Wasserstein Distance“ (SWD), posuzující distribuci obrazu.

---

<sup>1</sup>Dermatoskopie je neinvazivní diagnostická technika sloužící zejména k vyšetřování pigmentových kožních lézí. Použití dermoskopu, ručního optického zařízení složeného ze zvětšovací čočky a světelného zdroje, eliminuje odraz na povrchu pokožky a zlepšuje vizualizaci jejích hlubších úrovní. [65]

Na základě vizuálního Turingova testu a SWD lze konstatovat, že se Baur et al. úspěšně podařilo za pomoci progresivně rostoucího GAN syntetizovat vysoce realistické dermatoskopické obrazy kožních lézí. Ukázka těchto snímků je na obrázku 3.1.

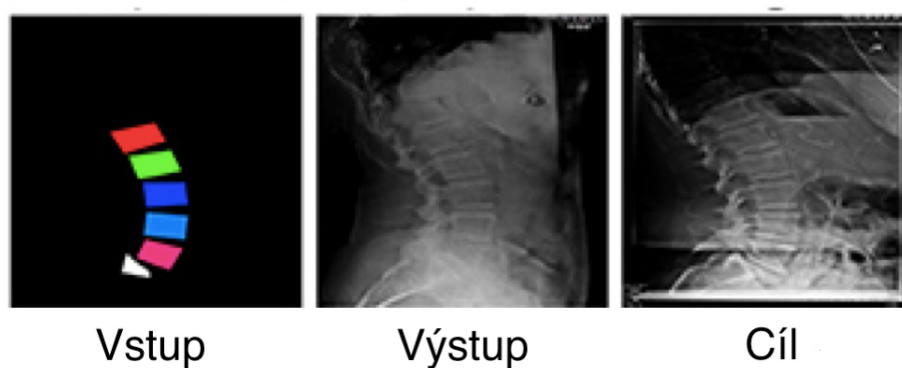


Obr. 3.1: Ukázka syntetických obrazů kožních lézí vygenerovaných modelem PGAN v publikaci [64].

Příkladem publikace zaměřené na podmíněnou syntézu medicínských obrazů je článek Galbusera et al. [67] zabývající se syntézou radiologických snímků páteře pro využití během *in silico* testování<sup>2</sup> nových medicínských technologií. Hlavním cílem práce bylo generovat syntetické RTG snímky na základě upravených sagitálních RTG snímků bederní páteře. Na snímcích byly manuálně vyznačeny čtyři rohy těl obratlů mezi L1 a L5 a čtyři body popisující horní část křížové kosti a na základě těchto souřadnic byly snímky transformovány na obraz, v němž jsou obratle reprezentovány barevně rozlišenými konvexními čtyřúhelníky. Na základě těchto snímků, zobrazujících přibližný tvar bederních obratlů a křížové destičky vygeneroval model GAN určený pro „image-to-image“ transformaci syntetické RTG obrazy páteře. Ukázka vstupního a výstupního obrazu je na obrázku 3.2.

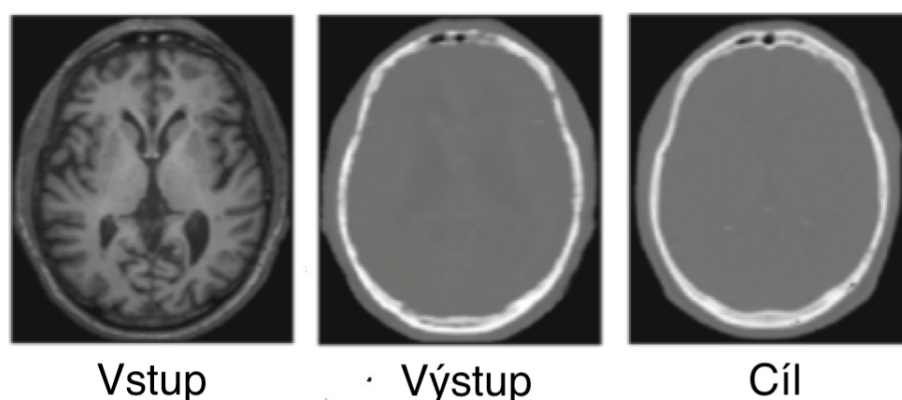
V publikaci Nie et al. [69] bylo představeno využití „3D-GAN“ pro generování CT obrazů, jež jsou důležité například při plánování radiační léčby nebo korekci útlumu v MRI nebo PET skeneru. Během pořizování CT snímku je pacient vystaven radiaci, která může poškodit zdravé buňky a zvýšit riziko rakoviny. Oproti tomu pořizování MRI snímků je bezpečnou zobrazovací technikou a pro snížení rizika pro

<sup>2</sup>Avicenna Alliance definuje *in silico* testování jako: „Použití individualizované počítačové simulace při vývoji nebo regulačním hodnocení léčivého přípravku, zdravotnického prostředku nebo lékařského zásahu [68].“



Obr. 3.2: Vstupní obraz s barevně vyznačenými obratli L1-L5 a horní částí křížové kosti byl transformován na výstupní obraz páteře. Pro srovnání lze využít skutečný RTG snímek páteře. Upraveno z [67].

pacienta je tedy žádoucí odhadovat CT snímky z MRI snímků. Navržená metoda je založena na „3D-GAN“, v němž je jako generátor použita plně konvoluční neuronová síť soutěžící s diskriminátorem. Kriteriační funkce generátoru je rozšířena o výraz pro rozdíl gradientu a obrazu „Gradient Difference Loss“ s cílem zachovat ostrost generovaného CT obrazu a je využit „Auto-Context Model“ pro zpřesnění výstupu generátoru. Hodnocení na dvou skutečných souborech CT/MRI dat ukázalo, že tato metoda dokáže efektivně transformovat MRI snímky na CT obrazy a svými výsledky metoda překonala tři nejmodernější metody, s nimiž byla porovnávána. Ukázka vstupního, výstupního a srovnávacího obrazu je na obrázku 3.3.



Obr. 3.3: Vstupní MRI snímek byl modelem 3D GAN transformován na syntetický CT snímek a porovnán s reálným CT snímkem. Upraveno z [69].

## 4 Experimentální data

V praktické části práce bylo využito několik datových sad, jež jsou popsány v následujících podkapitolách.

### 4.1 Dermatoskopické snímky kožních lézí

Pro generování obrazů kožních lézí byla použita data z „ISIC 2020 Challenge Dataset“ [70] a „ISIC 2017 Challenge Dataset“ [71], volně dostupná na portálu „ISIC Challenge“.

#### 4.1.1 ISIC 2020 Challenge Dataset

##### Popis databáze

Datová sada „ISIC 2020 Challenge Dataset“ obsahuje 33 126 dermatoskopických tréninkových snímků benigních a maligních kožních lézí od více než 2 000 pacientů. Část těchto snímků byla využita pro trénování nepodmíněného modelu DCGAN, jehož vstupem je kromě vektoru šumu pouze sada reálných snímků.

##### Předzpracování dat

Z datové sady bylo vybráno prvních 10 000 snímků formátu JPEG (z angl. „Joint Photographic Experts Group“), které byly v rámci předzpracování načteny jako tříkanálové obrázky datového typu „UInt8“ o rozměrech  $64 \times 64 \times 3$ . Ukázka předzpracovaných snímků je na obrázku 4.2. Po zpracování všech souborů byly obrázky převedeny do datové struktury zvané pole („array“) a uloženy v komprimovaném formátu „NPZ“. Výsledný soubor obsahuje multidimenzionální pole s rozměry  $10000 \times 64 \times 64 \times 3$ . Rozměry pole jsou definovány ve formátu  $B \times H \times W \times C$ , kde písmeno  $B$  udává počet snímků v datové sadě,  $H$  udává výšku snímků,  $W$  šířku a  $C$  počet kanálů.



Obr. 4.1: Ukázka předzpracovaných snímků z „ISIC 2020 Challenge Dataset“.

Stejným způsobem byla vytvořena také datová sada s rozlišením  $128 \times 128$  a  $256 \times 256$  pixelů. Před trénováním a testováním modelu byla vybraná datová sada načtena, dále byly hodnoty převedeny do běžně využívaného formátu „float32“ a normalizovány na interval  $\langle -1, 1 \rangle$ . V průběhu trénování modelu DCGAN byla využita augmentace snímků, konkrétně vertikálním a/nebo horizontálním otočením obrazu vždy s pravděpodobností 0.2.

## 4.1.2 ISIC 2017 Challenge Dataset

### Popis databáze

Datová sada „ISIC 2017 Challenge Dataset“ obsahuje trénovací, validační a testovací sadu dermatoskopických snímků benigních a maligních kožních lézí. Pro každý dermatoskopický snímek je zveřejněna také odpovídající binární maska vyznačující hranice léze a anotace dokumentující dermatoskopické rysy a klasifikaci léze. Binární masky byly vytvořeny odborným lékařem pomocí poloautomatického nebo manuálního procesu a spolu s dermatoskopickými snímky byly použity pro trénování a testování podmíněného modelu *pix2pix*.

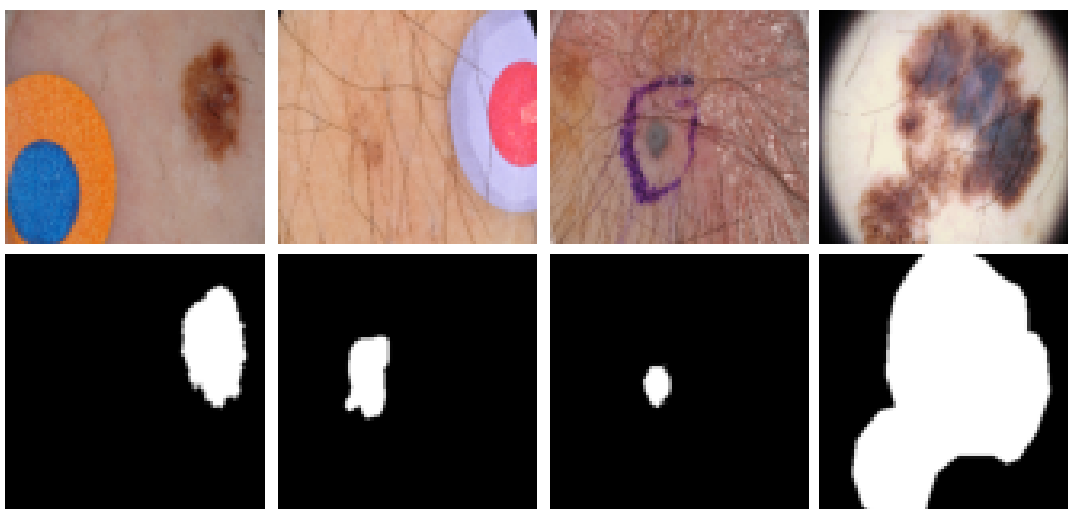
Trénovací sada obsahuje 2 000 snímků, konkrétně 375 snímků melanomu, 254 snímků seboroické keratózy a 1 372 snímků benigních név. Testovací sada zahrnuje 600 snímků lézí těchto tříd, ovšem bez známého podílu zastoupení. V datových sadách jsou dva typy souborů, a to dermatoskopické snímky ve formátu JPEG a odpovídající binární masky ve formátu jednokanálového 8bitového PNG (z angl. „Portable Network Graphics“). Masky je tedy tvořena pixely hodnot 0 nebo 255, kde pixely s hodnotou 0 reprezentují oblast mimo lézi a pixely s hodnotou 255 oblast uvnitř léze.

### Předzpracování dat

Předzpracování proběhlo obdobným způsobem jako u první datové sady. Soubory byly načteny jako tříkanálové obrázky datového typu „UInt8“ a velikosti  $256 \times 256 \times 3$ . Ukázka předzpracovaných snímků a masek je na obrázku 4.2. Výsledný soubor pro trénovací sadu obsahuje dvě multidimenzionální pole. První pole obsahuje předzpracované snímky a druhé předzpracované binární masky, přičemž obě pole mají rozměry  $2000 \times 256 \times 256 \times 3$ . Soubor pro trénovací sadu obsahuje dvě obdobné pole, ovšem s rozměry  $600 \times 256 \times 256 \times 3$ .

Před trénováním a testováním modelu byly datové sady opět načteny a normalizovány na interval  $\langle -1, 1 \rangle$ .





Obr. 4.2: Ukázka předzpracovaných snímků a masek z „ISIC 2017 Challenge Dataset“.

## 4.2 Snímky CT páteře

### 4.2.1 Popis databáze

Dále byly v praktické části využity CT snímky páteře 12 anonymních pacientů, poskytnuté vedoucím práce. U 10 pacientů se jednalo o abdominální snímky zachycující celou páteř a u zbylých 2 pacientů o snímky pouze lumbálně–sakrální části páteře. U pacientů se vyskytují různé patologické stavy páteře a obratlů včetně osteoporézy, zborcení obratle, osteolytických a osteoblastických lézí a dalších.

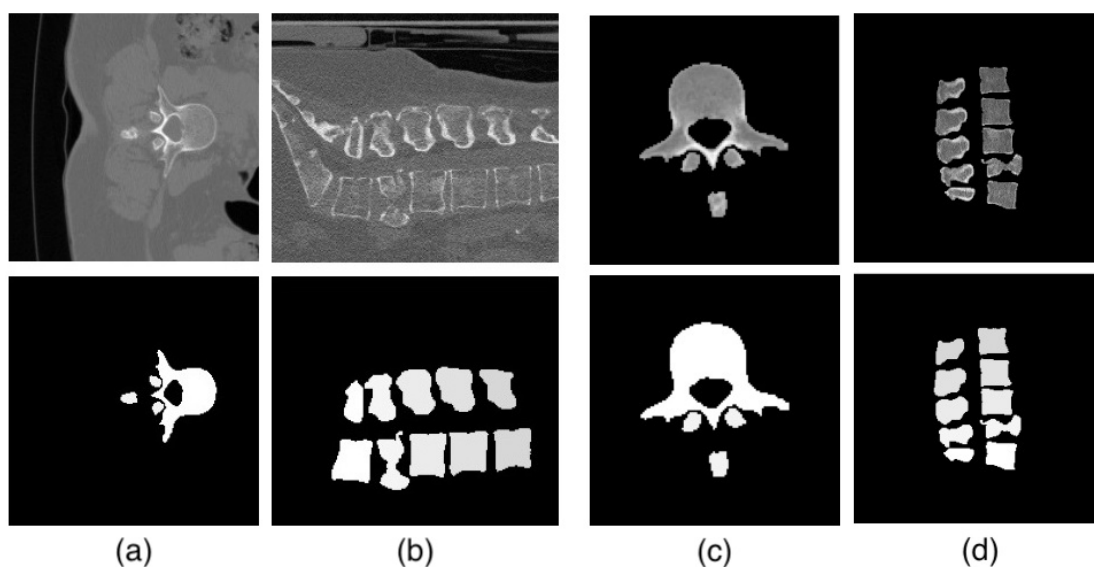
Ke každému pacientovi se vztahují 2 soubory formátu NIfTI (z angl. „Neuroimaging Informatics Technology Initiative“). První soubor obsahuje CT data uložena v 32bitových hodnotách vyjadřujících Hounsfieldovy jednotky (HU) bez nutnosti lineární transformace. V druhém souboru je uložena binární maska vyznačující segmentaci páteře v těchto datech. Pro účely této práce byly vytvořeny sady axiálních a sagitálních řezů, které byly dále předzpracovány.

### 4.2.2 Předzpracování dat

Prvním krokem předzpracování snímků byla jejich normalizace. Hodnoty CT snímku v užitečném rozsahu  $-1024 - 1976$  HU byly přičtením hodnoty 1024 převedeny do kladných hodnot a bylo provedeno prahování se spodním prahem 0 HU a horním prahem 3000 HU. Hodnoty byly dále převedeny na datový typ „float32“ a normalizovány na interval  $(0, 1)$ . V dalším kroku byla binární maska násobením aplikována

na CT snímek a z výsledného snímku byl vyříznut čtyřúhelník obsahující oblast nenulových hodnot.

V tomto bodě se mírně rozchází předzpracování axiálních a sagitálních snímků. V případně axiálních snímků byly z datové sady automaticky vyřazeny všechny řezy obsahující pouze nulové hodnoty a při ořezu byla vždy ponechána celá oblast nenulových hodnot. Při zpracování sagitálních snímků byly řezy vybrány manuálně a ořezána byla pouze oblast lumbální páteře, snímaná u všech pacientů. Axiální snímky byly poté přidáním nulových hodnot na okraje vyplněny na velikost  $128 \times 128$  pixelů a sagitální na velikost  $256 \times 256$  pixelů. Stejný proces ořezu a vyplnění byl proveden i s maskami a všechny snímky byly rotovány o 90 stupňů. Ukázka CT snímků a binárních masek před a po předzpracování je na obrázku 4.3.



Obr. 4.3: Ukázka axiálního (a) a sagitálního (b) řezu před předzpracováním. Axiální (c) a sagitální (d) snímek a maska po předzpracování.

Takto předzpracované snímky byly rozděleny v poměru 3:1 do tréninkové a testovací sady. Testovací sada obsahuje snímky 3 pacientů vybraných tak, aby se mezi testovacími snímky nacházely fyziologické i patologické tvary obratlů zajímavé pro analýzu a diskuzi. Výsledné tréninkové datové sady mají rozměry  $3330 \times 128 \times 128 \times 1$  pro axiální snímky a  $389 \times 256 \times 256 \times 1$  pro sagitální snímky. Tyto datové sady jsou uloženy a načítány stejným způsobem jako sady snímků kožních lézí.

## 5 Navržené modely pro generování medicínských obrazů

V teoretickém úvodu byla představena celá řada modelů GAN schopných syntézy kvalitních medicínských obrazů. Dvěma základními úlohami v této oblasti jsou nepodmíněná a podmíněná syntéza obrazu, jež jsou předmětem praktické části této práce.

Pro zkoumání nepodmíněného generování medicínských obrazů byl vybrán hluboce konvoluční GAN schopný syntézy obrazu pouze na základě vstupního vektoru šumu. DCGAN byl trénován na datové sadě dermatoskopických snímků z „ISIC 2020 Challenge Dataset“, popsané v 4.1.1, s cílem generovat realistické obrazy kožních lézí. Pro splnění tohoto cíle byl navržen také model *pix2pix*, který na rozdíl od předchozího modelu syntetizuje obraz na základě obrazového vstupu, čímž provádí podmíněnou syntézu. Pro tuto aplikaci byl model trénován a testován na datových sadách pocházejících z „ISIC 2017 Challenge Dataset“, popsanych v 4.1.2.

Další aplikací modelu *pix2pix* umožňující zkoumání podmíněné syntézy je generování obrazů páteře na základě binární masky. Pro trénování a testování modelu jsou použity datové sady popsané v kapitole 4.2, které obsahují axiální a sagitální CT snímky páteře spolu s binárními maskami.

V následujících kapitolách je blíže popsána architektura a proces učení těchto modelů.

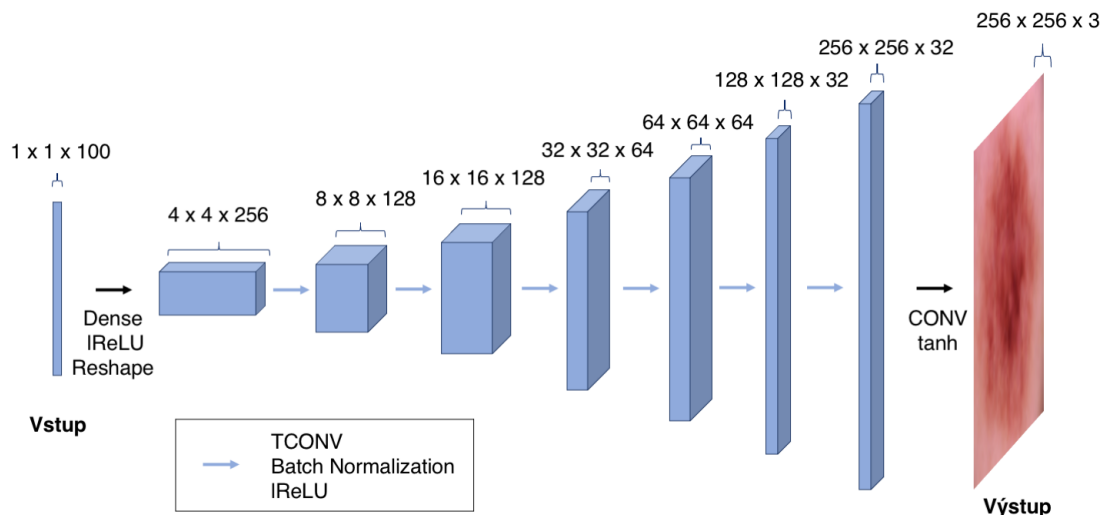
### 5.1 Hluboce konvoluční GAN

Hluboce konvoluční GAN, poprvé popsáný v publikaci Radford et al. [31], byl prvním významným krokem ke stabilizaci učícího procesu GAN. Rozsáhlým zkoumáním modelu byla identifikována architektura umožňující stabilní trénink generativních modelů s vyšším rozlišením a hlubokými vrstvami [31], kterou lze využít pro generování jednoduchých medicínských obrazů. Model DCGAN se skládá ze dvou hluboce konvolučních neuronových sítí, jedné zastávající roli generátoru a druhé v roli diskriminátoru.

#### 5.1.1 Generativní síť

Vstupem generátoru je vektor náhodných hodnot šumu z normální distribuce dat. Tento vektor je vrstvou „Reshape“ transformován na matici, jež je sérií bloků nadzorována až na výstupní rozměry. Základem každého bloku je vrstva transponované konvoluce provádějící inverzní konvoluční operaci, během níž dochází k transformaci

hrubých charakteristických rysů na detailnější výstup. Výstup vrstvy je normalizován vrstvou *Batch Normalization* (kromě výstupu poslední vrstvy) a aktivován funkcí lReLU. Hyperparametry vrstev jsou zvoleny tak, aby při průchodu blokem došlo ke dvojnásobnému zvětšení výšky  $H$  a šířky  $W$  matice a zároveň byl progresivně snižován počet kanálů obrazu  $C$ . Počet filtrů poslední konvoluční vrstvy definuje počet kanálů ve výstupním obraze. Schéma sítě generátoru je na obrázku 5.1.

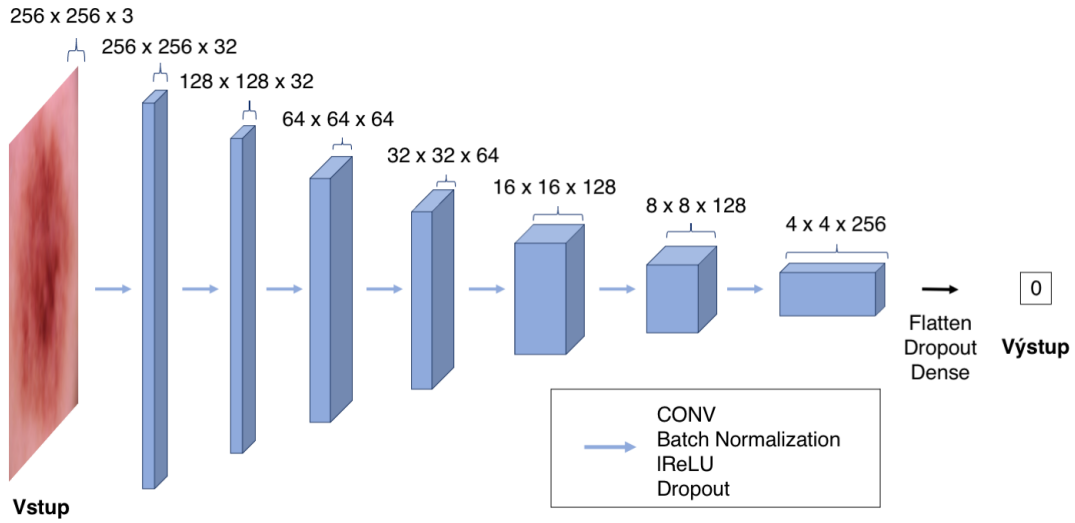


Obr. 5.1: Schéma architektury generativní sítě. Vstupem je vektor šumu s rozměry  $1 \times 1 \times 100$ , jež je 6 bloky transponované konvoluce transformován na výstupní obraz s rozměry  $256 \times 256 \times 3$ .

### 5.1.2 Diskriminativní síť

Vstupem diskriminativní sítě je obraz vygenerovaný generativní sítí nebo snímek z množiny tréninkových dat. Z obrazové matice jsou sériově spojenými konvolučními bloky extrahovány příznaky. Konvoluční vrstvy zároveň provádí i podvzorkování matice, čímž nahrazují klasickou podvzorkovací vrstvu. Výstupy konvolučních vrstev jsou normalizovány (kromě první), aktivovány funkcí lReLU a některé hodnoty jsou náhodně vynulovány vrstvou *Dropout*. Díky vynechání *Batch Normalization* vrstvy u poslední vrstvy generátoru a první vrstvy diskriminátoru jsou sítě schopny naučit se střední hodnotu a rozptyl distribuce dat. V síti diskriminátoru jsou progresivně zvyšovány počty příznakových map, které jsou na závěr zpracovány poslední plně propojenou vrstvou se sigmoidální aktivací. Výstupem diskriminátoru je jedna hodnota představující pravděpodobnost, že daný vstup náleží do množiny

tréninkových dat. Tato hodnota slouží pro výpočet kritériální funkce a optimalizaci modelu. Schéma diskriminativní sítě je znázorněno na obrázku 5.2.



Obr. 5.2: Schéma architektury diskriminativní sítě, jež ze vstupu o rozměrech  $256 \times 256 \times 3$  extrahuje 7 konvolučními bloky příznaky a na závěr vyhodnocuje příslušnost vstupu do množiny tréninkových dat.

### 5.1.3 Učení a optimalizace modelu

Cíl učícího procesu DCGAN je matematicky vyjádřen vzorcem 2.1. Ověřeným postupem při učení modelů GAN je vytvoření samostatných *mini-batch* reálných snímků a generovaných snímků, jež jsou diskriminátorem klasifikovány zvlášť. Reálným snímkům je přiřazena cílová hodnota 1, syntetickým snímkům hodnota 0 a jako kritériální funkce je využita binární křížová entropie (BCE), vhodná pro binární klasifikaci. Matematický vzorec pro výpočet BCE *mini-batch* je:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{n=1}^N [y \log(\hat{y}_n) + (1 - y) \log(1 - \hat{y}_n)], \quad (5.1)$$

kde  $y$  je cílová hodnota („ground truth“),  $\hat{y}_n$  je predikovaná hodnota a  $N$  vyjadřuje počet snímků v *mini-batch* [72].

Optimalizace modelu probíhá metodou adaptivního odhadu momentu (Adam).

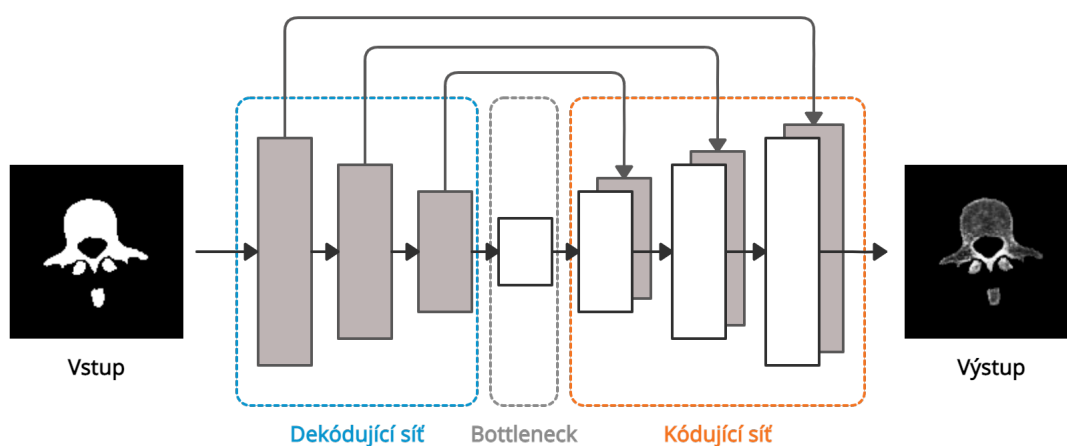
I navzdory stabilizačním prvkům mohou drobné změny architektury a hyperparametrů vést k selhání modelu. Při odchýlení od publikované architektury je proto třeba dbát na vyvážení výkonu diskriminátoru a generátoru, jež je podmínkou stabilního tréninku.

## 5.2 Model „pix2pix“

Návrh modelu *pix2pix* vychází ze studie publikované Isola et al. [33], v níž byla tato modifikace podmíněného modelu cGAN poprvé představena. Od té doby byl *pix2pix* úspěšně aplikován na řadu zadání z oblasti medicíny zahrnujících „Image-to-image“ transformaci. Model využívá generativní síť založenou na architektuře „U-Net“, diskriminativní síť typu „PatchGAN“ a způsob optimalizace podobný modelu cGAN.

### 5.2.1 Generativní síť typu „U-Net“

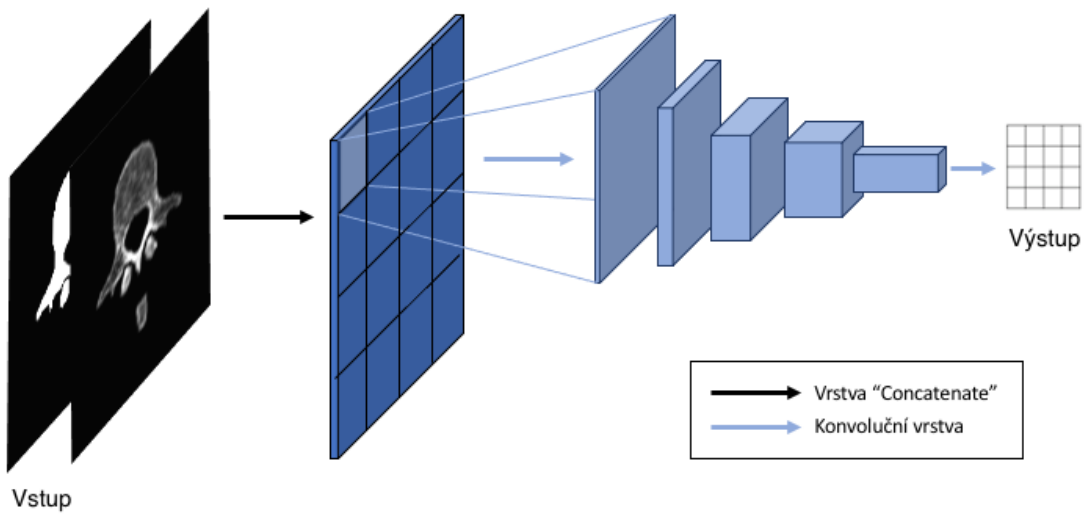
Generátorem modelu *pix2pix* je konvoluční síť s architekturou „U-Net“, původně navrženou pro segmentaci biomedicínských dat [73]. První část generátoru je typická konvoluční síť extrahující příznaky obrazu konvolučními filtry a redukující prostorové rozměry obrazové matice pomocí podvzorkování. Vrstva, v níž jsou příznakové mapy zredukovány až na velikost  $1 \times 1$ , se označuje jako „bottleneck“. Na ni navazuje expanzivní část sítě, kombinující příznaky procházející *bottleneck* vrstvou s informacemi z vrstev extrakční části. Spojení dekódující a kódující části sítě, ilustrované na obrázku 5.3, umožňuje lepší šíření kontextových informací do vrstev s vyšším rozlišením. Takovýto generátor je schopen na základě masky generovat vysoce kvalitní medicínské obrazy s vysokým rozlišením.



Obr. 5.3: Vstupem generátoru je binární maska, z níž jsou dekódující částí sítě extrahovány příznaky. V druhé části sítě jsou příznaky a informace z dekódující části zakomponovány do nově vzniklého obrazu.

### 5.2.2 Diskriminativní síť typu „PatchGAN“

Diskriminátor „PatchGAN“ vychází z architektury hluboce konvolučních sítí. Vstupem této sítě je binární maska spojená vrstvou „Concatenate“ s obrazovým výstupem generátoru. Reálnost obrazu je na rozdíl od diskriminátoru použitého v modelu DCGAN klasifikována na základě lokálních oblastí (angl. „patch“). Reálnost překrývajících se oblastí je konvolučními vrstvami klasifikována samostatně a výsledná matice hodnot pravděpodobnosti je při trénování modelu porovnávána s očekávaným výstupem. Diskriminativní síť je ilustrována na obrázku 5.4.



Obr. 5.4: První vrstva diskriminativní sítě spojuje binární masku s reálným snímkem nebo výstupem generátoru, z nichž jsou sérií konvolučních vrstev extrahovány příznaky. Každé lokální oblasti obrazu je poslední konvoluční vrstvou se sigmoidální aktivací určena míra pravděpodobnosti, že náleží do reálné distribuce dat.

### 5.2.3 Učení a optimalizace modelu

Základní učící cíle modelu *pix2pix* zůstávají stejné jako u ostatních modelů GAN. Generátor se snaží generovat výstupy, jež diskriminátor nedokáže rozlišit od reálných dat, zatímco diskriminátor je trénován v rozpoznávání syntetických obrazů. Tento cíl je vyjádřen kritériální funkcí podmíněného modelu cGAN, popsanou vzorcem:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))], \quad (5.2)$$

kde  $x$  představuje vstup, na kterém zakládáme generativní proces (v tomto modelu binární masku),  $z$  značí šum a  $y$  předpokládaný výstup (reálný snímek). Výraz  $D(x, y)$  pak vyjadřuje výstup diskriminátoru pro binární masku a reálný snímek. Výstup diskriminátoru pro binární masku a syntetický snímek  $G(x, z)$  je ve vzorci reprezentovaný výrazem  $D(x, G(x, z))$ . Diskriminátor se snaží tuto funkci maximalizovat, zatímco generátor je trénován, aby ji minimalizoval.

V případě tohoto modelu však nestačí, aby generátor generoval dostatečně kvalitní obrazy, které diskriminátor klasifikuje jako reálné. Dalším cílem generátoru je, aby byly obrazy totožné s odpovídajícími reálnými snímky, což vyjadřuje funkce označovaná jako „nejmenší absolutní odchylky“ (angl. „Least Absolute Deviations“) nebo také „L1 vzdálenost“. Tato kritériální funkce, popsaná vzorcem 5.3, hodnotí sumu absolutních rozdílů mezi předpokládaným výstupem generátoru  $y$  a jeho skutečným výstupem  $G(x, z)$ .

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1] \quad (5.3)$$

Celý učicí proces modelu lze vyjádřit vzorcem 5.4, jenž zahrnuje obě kritériální funkce definované výše. Jedná se o váhovaný součet těchto funkcí, přičemž „L1 vzdálenost“ je násobena koeficientem  $\lambda$ .

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (5.4)$$

Vzorce vycházejí z publikace Isola et al. [33].



## 6 Implementace algoritmů

Veškeré programové řešení bylo implementováno v programovacím jazyce Python, konkrétně ve verzi Python 3.8.5. Pro tvorbu generativních soutěživých sítí byla využita knihovna pro hluboké učení Tensorflow s pomocí vysokoúrovňového aplikačního rozhraní (API) Keras. Keras je volně dostupná knihovna navržená pro experimenty s hlubokými neuronovými sítěmi, jež umožňuje uživatelsky přívětivé a efektivní využití Tensorflow. Vzhledem k výpočetní náročnosti algoritmů probíhala optimalizace prostřednictvím grafické procesorové jednotky NVIDIA GeForce GTX 1060 6GB.

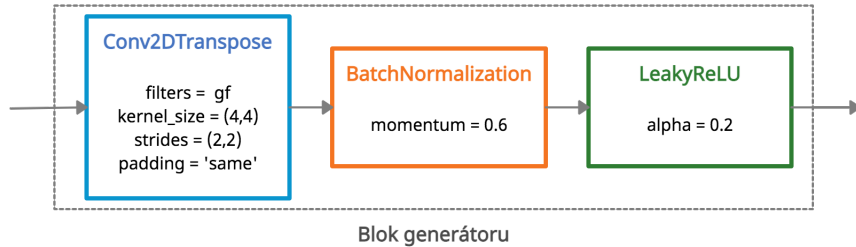
### 6.1 Hluboce konvoluční GAN

Implementace hluboce konvolučního GAN probíhala po částech. V první řadě byl v programovém prostředí realizován model pro generování výstupů s rozměry  $64 \times 64 \times 3$ , vycházející z publikace Radford et al. [31]. Po změnách architektury a hyperparametrů bylo dosaženo uspokojivých výsledků, tj. natrénovaný generátor generoval barevné obrazy s rozlišením  $64 \times 64$  s uspokojivou kvalitou. Dalším krokem byl přechod na rozlišení  $128 \times 128$ , jež vyžadoval přidání bloků do architektury generátoru a diskriminátoru a drobné změny vybraných hyperparametrů. Při návrhu sítě s větším rozlišením výstupů bylo důležité zejména využití prvků oslabujících síť diskriminátoru tak, aby mohl probíhat stabilní trénink modelu. Stejným způsobem byl proveden ještě jeden přechod na vyšší rozlišení a výsledkem implementace je hluboce konvoluční GAN generující barevné obrazy s rozlišením  $256 \times 256$ , popsány v následujících podkapitolách.

#### 6.1.1 Generativní síť

Do první plně propojené **Dense** vrstvy generátoru s 4096 neurony vstupuje vektor 100 náhodných hodnot z Gaussovy distribuce dat. Následuje aktivační vrstva **LeakyReLU** a vektor je vrstvou **Reshape** tvarován na matici s rozměry  $4 \times 4 \times 256$ . Matice je bloky generátoru nadzvorkována na výstupní obraz s rozměry  $256 \times 256 \times 3$ . Pro zvýšení rozlišení matice se nabízejí dva postupy. Jednoduchým řešením je použití vrstvy **UpSampling2D**, jež ve výstupu opakuje řádky a sloupce poskytnuté jako vstup. Samotná vrstva **UpSampling2D** však není schopna doplnit do výstupní obrazové matice užitečné informace a musí být následována vrstvou **Conv2D**, která se učí tento výstup smysluplně interpretovat. Kombinaci těchto dvou vrstev lze nahradit vrstvou transponované konvoluce **Conv2DTranspose**, jež provádí inverzní konvoluční operaci, čímž zvyšuje rozlišení obrazu a zároveň vstup smysluplně transformuje. Lepších výsledků bylo dosaženo při použití vrstvy **Conv2DTranspose** s hyperparametrem

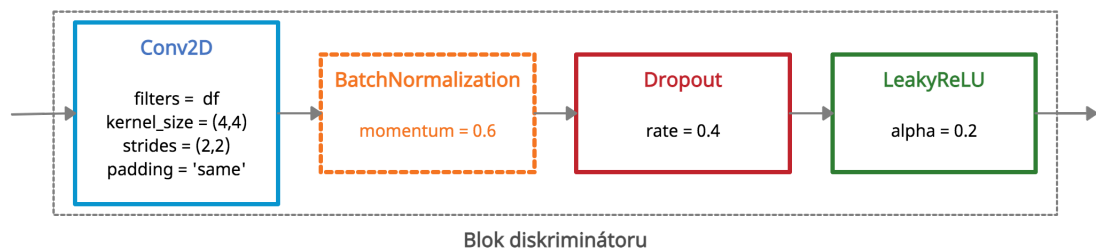
$\text{stride} = (2, 2)$ , díky čemuž byly tyto vrstvy použity jako hlavní prvky generátoru. Počet konvolučních filtrů  $gf$  progresivně roste z 32 na 256. Po průchodu sérií bloků transponované konvoluce, jež jsou popsány na obrázku 6.1, je poslední konvoluční vrstvou **Conv2D** snížen počet kanálů na 3 a hodnoty výstupní obrazové matice jsou aktivací funkcí **tanh** normalizovány na interval  $\langle -1, 1 \rangle$ . Tímto způsobem je vygenerován syntetický obraz, který v tréninkové fázi slouží jako vstup diskriminátoru.



Obr. 6.1: Blok transponované konvoluce využitý v generátoru modelu DCGAN. Generátor implementované sítě obsahuje 6 takovýchto bloků. Jediným měnícím se hyperparametrem je počet konvolučních filtrů, definovaný pomocnou proměnnou  $gf$ .

### 6.1.2 Diskriminativní síť

Do diskriminátoru vstupuje reálný snímek z „ISIC 2020 Challenge Dataset“ nebo výstup generátoru, vždy ve formě matice s rozměry  $256 \times 256 \times 3$ . Konvoluční vrstvy **Conv2D** mají obdobné hyperparametry jako vrstvy transponované konvoluce v generátoru, počet jejich konvolučních filtrů však progresivně klesá z 256 na 32. S výjimkou první konvoluční vrstvy jsou následovány vrstvou **Batch Normalization**. Další prvky bloků diskriminátoru jsou popsány na obrázku 6.2.



Obr. 6.2: Konvoluční blok využitý v diskriminátoru modelu DCGAN. V diskriminátoru implementovaného modelu se tento blok opakuje 7krát, s různým počtem konvolučních filtrů  $df$ .

Na sérii konvolučních bloků navazuje klasifikační část diskriminátoru. Výstup poslední konvoluční vrstvy (matice s rozměry  $4 \times 4 \times 256$ ) je vrstvou `Flatten` transformován na vektor 4096 hodnot. Následně je aplikována vrstva `Dropout` a plně propojenou vrstvou `Dense` s 1 neuronem a aktivační funkcí `sigmoid` je vyhodnocena pravděpodobnost, s níž vstup náleží do tréninkové sady dat.

### 6.1.3 Učení a optimalizace modelu

Parametry vrstev jsou na počátku tréninku nastaveny na malé náhodné hodnoty z Gaussovy distribuce. Inicializace je implementována vrstvou `RandomNormal` s průměrem 0 a směrodatnou odchylkou 0,02 a váhy jsou během tréninku optimalizovány metodou Adam s parametry:  $\alpha = 0,0002$ ,  $\beta_1 = 0,5$  a  $\beta_2 = 0,999$ .

V každé iteraci tréninku je z datové sady náhodně vybrána *mini-batch* 16 reálných snímků, generátorem je vytvořena *mini-batch* 16 syntetických obrazů a zvlášť jsou klasifikovány diskriminátorem. Pro klasifikaci do dvou tříd (1 pro reálná a 0 pro syntetická data) je použita kriteriální funkce `BinaryCrossentropy`. Pro oslabení diskriminátoru je navíc hyperparametr `label_smoothing` nastaven na hodnotu 0,1. Diskriminátor tudíž u reálných snímků očekává hodnotu 0,9 místo 1 a u syntetických 0,1 místo 0, čímž je snížena jeho klasifikační schopnost a podpořena stabilita tréninku.

## 6.2 Model „pix2pix“

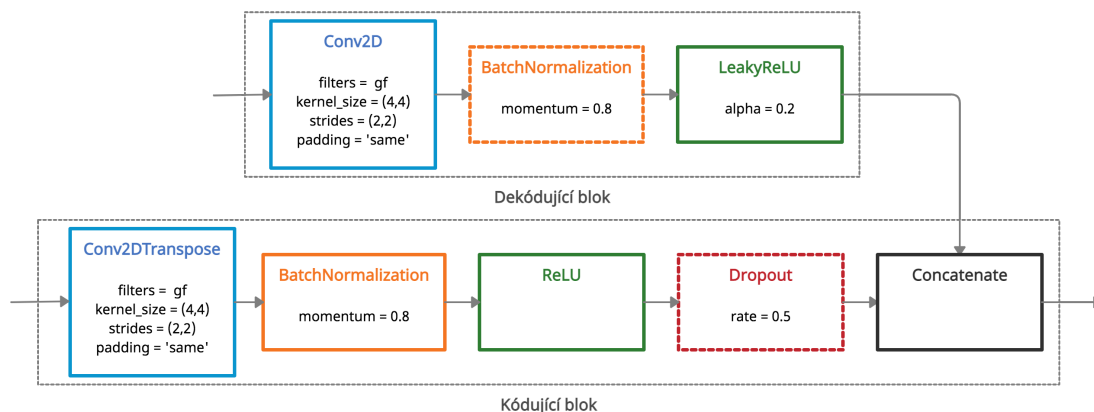
Model *pix2pix* byl v praktické části aplikován na 3 různé datové sady. Každá datová sada obsahuje vzorky jiných rozměrů, čemuž musí být přizpůsobena i architektura sítí. I přes to však implementace jednotlivých modelů obsahují společné prvky.

### 6.2.1 Generativní síť

Vstupem generátoru je binární maska, jež je průchodem sítí transformována na výstupní syntetický obraz. Dekódující část generátoru se skládá z několika sériově navazujících bloků. Každý blok začíná konvoluční vrstvou `Conv2D` extrahující příznaky a zároveň sloužící i k podvzorkování obrazové matice pomocí parametru `stride`. Ve všech blocích kromě prvního jsou výstupy konvoluční vrstvy normalizovány `BatchNormalization` a každý blok je zakončen aktivační vrstvou `LeakyReLU`.

Vrstva *bottleneck* je implementována jako konvoluční vrstva `Conv2D` s aktivační vrstvou `ReLU` a hyperparametry `kernel_size = (4,4)`, `strides = (2,2)` a `padding = 'same'`.

Bloky tvořící kódující část obsahují vrstvy **Conv2DTranspose**, jejichž výstupy jsou normalizovány vrstvami **BatchNormalization**. Další vrstvou je aktivace typu **ReLU** v prvních několika blocích následovaná vrstvou **Dropout**. Vrstvy **Dropout** jsou hlavním zdrojem stochasticity v tomto modelu a nahrazují tak roli vektoru šumu, který je vstupem nepodmíněných modelů jako je DCGAN. Kódující blok je zakončen vrstvou **Concatenate**, jež spojuje výstup bloku s výstupem příslušného dekódujícího bloku přes všechny kanály. Hyperparametry a sled jednotlivých vrstev je naznačen na obrázku 6.4.

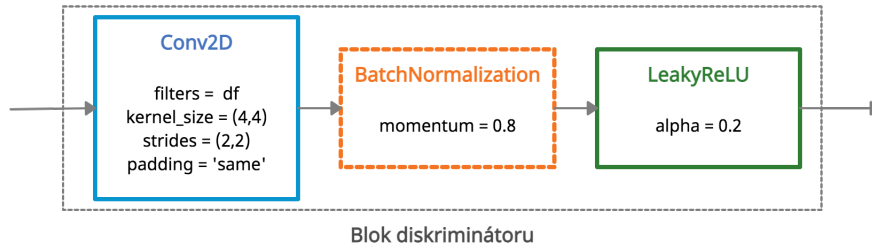


Obr. 6.3: Ukázka vrstev dekódujícího a kódujícího bloku s proměnným počtem konvolučních filtrů  $gf$ .

Počet filtrů v konvolučních vrstvách dekódující části je postupně zvyšován až po *bottleneck*, od něž se stejným způsobem počet filtrů ve vrstvách transponované konvoluce snižuje. Vrstvou **Concatenate** jsou spolu vždy spojeny vrstvy se stejným počtem filtrů a stejnými rozměry výstupu. Výstupem poslední vrstvy **Conv2DTranspose** s aktivací **tanh** je obrazová matice s hodnotami v intervalu  $(-1, 1)$  a stejnými rozměry jako vstupní maska.

## 6.2.2 Diskriminativní síť

Implementace „PatchGAN“ se od diskriminátoru modelu DCGAN liší jen drobnými změnami. Před vstupem do prvního konvolučního bloku je vrstvou **Concatenate** spojena binární maska a k ní odpovídající reálný snímek nebo syntetický obraz. Poté jsou z obrazové matice konvolučními bloky extrahovány příznaky jako v běžném diskriminátoru. Rozdíl nastává při klasifikaci výstupů, kterou v případě „PatchGAN“ provádí poslední **Conv2D** vrstva se sigmoidální aktivací. Výstupem této vrstvy je dvourozměrná matice hodnot pravděpodobnosti pro jednotlivé oblasti vstupního obrazu, jež je využita pro optimalizaci modelu.



Obr. 6.4: Ukázka bloku diskriminátoru, v němž je počet konvolučních filtrů definovaný pomocnou proměnnou  $df$ .

### 6.2.3 Učení a optimalizace modelu

Optimalizace modelu je opět započata inicializací vah vrstvou `RandomNormal` a optimalizace dále pokračuje metodou Adam s parametry:  $\alpha = 0,0002$ ,  $\beta_1 = 0,5$  a  $\beta_2 = 0,999$ .

Trénink modelů *pix2pix* probíhal po *mini-batch* velikosti 4, čili v každé tréninkové iteraci byly na základě masky vygenerovány 4 syntetické obrazy a z množiny tréninkových dat byly náhodně vybrány 4 reálné snímky. Malé trénovací *mini-batch* vnášejí do učícího procesu další šum a mají tak regularizační efekt. Model byl testován s různými velikostmi *mini-batch*. Při zvolení velikosti 1, jak je navrhováno v původní publikaci, byly výsledky srovnatelné, naopak při použití větších *mini-batch* se kvalita výstupů snižovala. Kvůli urychlení tréninku byla tedy u všech modelů zvolena velikost *mini-batch* 4.

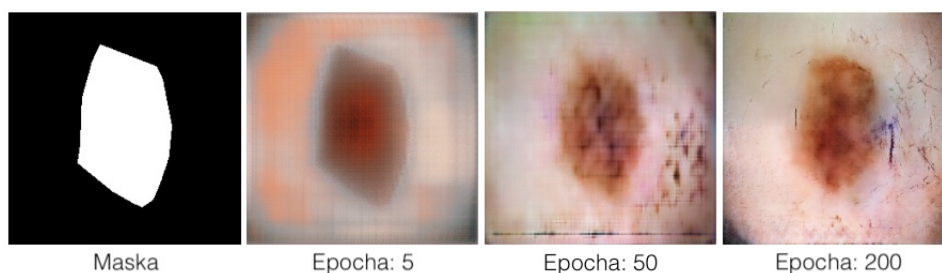
Základní kriteriální funkce modelu *pix2pix*, definovaná vzorcem 5.2, byla implementována funkcí `BinaryCrossentropy` a jako druhá kriteriální funkce generátoru byla využita funkce MEA (angl. „Mean Absolute Error“) váhovaná koeficientem  $\lambda = 100$ . Byla též testována kombinace kriteriálních funkcí MSE (angl. „Mean Squared Error“) a MEA, jež však dosahovala horších výsledků, a tudíž byla použita `BinaryCrossentropy`. Během trénování diskriminativní sítě je využita pouze `BinaryCrossentropy` a efekt této kriteriální funkce je váhováním hodnotou 0,5 oslaben na polovinu, čímž je podpořen trénink generátoru a celková stabilita učícího procesu.

Ani u modelu *pix2pix* však kriteriální funkce nijak nevypovídají o kvalitě výstupů modelu, a je proto náročné objektivně posoudit v jaké fázi je třeba ukončit trénink modelu. Pro sledování tréninku se tedy využívá subjektivní hodnocení výstupních obrazů a na základě kvalitativního vyhodnocení je zvolena verze modelu s nejlepšími výsledky.

### 6.2.4 Model „pix2pix“ pro generování obrazů kožních lézí

Specifikem implementace modelu „pix2pix“ pro generování obrazů kožních lézí jsou rozměry vstupní binární masky a reálných snímků. U tohoto modelu pracujeme se vzorky s rozměry  $256 \times 256 \times 3$ , čemuž musí odpovídat také architektura sítě generátoru a diskriminátoru. Dekódovací část generátorů je složena ze 7 bloků, u nichž stoupá počet konvolučních filtrů z 64 na 512. Rozměry matice při výstupu z vrstvy *bottleneck* jsou  $1 \times 1 \times 512$ . V kódovací části u 7 kódujících bloků ekvivalentně počty konvolučních filtrů klesají z 512 na 64 v posledním bloku. Poslední blok je následován vrstvou *Convolution2DTranspose*, která matici transformuje na výstupní rozměry  $256 \times 256 \times 3$ . Síť diskriminátoru je implementována 5 konvolučními bloky opět se stoupajícím počtem filtrů z 64 na 512. Výstupem této sítě je matice s rozměry  $16 \times 16 \times 1$ , jež je při tréninku porovnávána s předpokládanými hodnotami a na základě rozdílů mezi maticemi jsou upraveny váhy modelu.

Takto definovaný model je trénován způsobem popsáním v předchozí podkapitole. Na obrázku 6.5 je naznačen pokrok generátoru během učícího procesu. Na počátku generuje pouze obecné tvary bez ostrých hran, po 50 epochách učení již můžeme na obraze rozeznat kožní lézi a v průběhu dalších 150 epoch se generátor učí detailnější charakteristiky léze. Zhruba od učící epochy 200 začal trénink stagnovat a během dalších 100 epoch se u modelu neprojevovalo zlepšení.



Obr. 6.5: Pokrok generátoru během tréninkové fáze.

### 6.2.5 Model „pix2pix“ pro generování snímků páteře

Pro generování obrazů páteře byly využity dvě datové sady, každá o jiných rozměrech. Nejdříve bude popsána implementace modelu pro generování axiálních snímků.

Vstupní data tohoto modelu mají rozměry  $128 \times 128 \times 1$  a bylo tak nutné model mírně modifikovat. Počet kanálů vstupních a výstupních vrstev generátoru a diskriminátoru je závislý na rozměrech vstupních dat a není tedy třeba manuálně měnit hyperparametry sítě. Hlavními změnami v modelu je tím pádem snížení počtu bloků generátoru o 1, zbytek implementace je totožný.

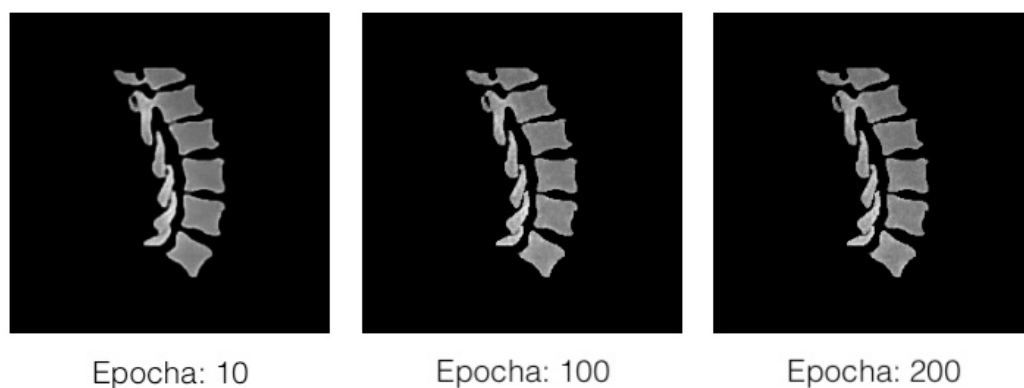
Do generátoru vstupuje obrazová matice s rozměry  $128 \times 128 \times 1$ , jež je 6 dekodujícími bloky podvzorkována až na výstupní rozměry vrstvy *bottleneck*  $1 \times 1 \times 512$ . Poté je proces obrácen a 6 kódujícími bloky je obrazová matice transformována zpět na rozměry  $128 \times 128 \times 1$ . Tento výstup je vrstvou **Concatenate** spojen s binární maskou a vstupuje do diskriminátoru. Výstupem diskriminátoru je matice s rozměry  $8 \times 8 \times 1$ , jež reprezentuje pravděpodobnosti pro lokální oblasti vstupního obrazu.

Výstupy generátoru v průběhu tréninku jsou srovnány na obrázku 6.6. S postupem učicího procesu je generátor schopen generovat detailnější výstupní obraz.



Obr. 6.6: Tréninkový proces modelu pro generování axiálních snímků páteře.

Implementace modelu pro generování axiálních snímků s rozměry  $256 \times 256 \times 1$  je obdobná jako u modelu pro generování obrazů kožních lézí, pouze s rozdílným počtem kanálů. Učící proces tohoto modelu je zaznamenán na obrázku 6.7.



Obr. 6.7: Tréninkový proces modelu pro generování sagitálních snímků páteře.

## 7 Diskuze dosažených výsledků

V následující části práce jsou prezentovány a diskutovány výsledky generování medicínských obrazů pomocí modelů GAN. Modely GAN lze hodnotit kvalitativními i kvantitativními metodami založenými na hodnocení kvality a rozmanitosti generovaných obrazů. Během kvalitativního hodnocení jsou generované obrazy subjektivně hodnoceny pozorovatelem, často ve srovnání s reálnými daty. Kvantitativní hodnocení GAN, jež je nejběžněji realizováno klasifikační neuronovou sítí [76], je již nad rámec této práce. Vizuální hodnocení výstupů hrálo důležitou roli i při trénování modelů. Během tréninkové fáze byly aktuální parametry modelu generátoru ve zvolených intervalech ukládány a následně byla pro každý model vybrána verze generátoru s nejvyšší kvalitou výstupního obrazu, jež byla použita pro testování.

### 7.1 Hluboce konvoluční GAN

Prvním modelem implementovaným v praktické části je hluboce konvoluční GAN aplikovaný na datovou sadu dermatoskopických snímků kožních lézí. Po uplynutí 645 epoch učení byla generativní síť modelu DCGAN využita pro generování syntetických obrazů s rozměry  $256 \times 256 \times 3$  z vektoru šumu.



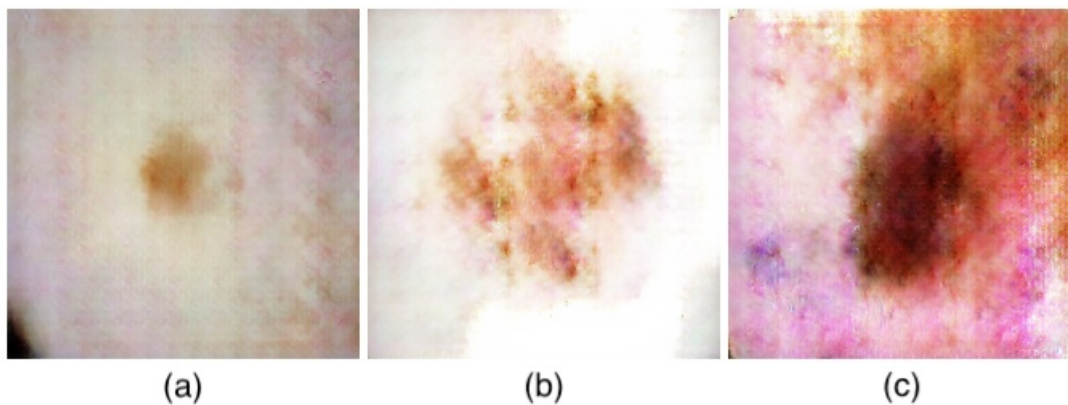
Obr. 7.1: Ukázka syntetických obrazů kožních lézí vygenerovaných modelem DCGAN.

Obrazy kožních lézí generované tímto modelem kopírují základní rysy dermatoskopických snímků. Z ukázky na obrázku 7.1 lze usoudit, že model generuje obrazy kožních lézí s charakteristickými barvami, siluetou a pozicí v rámci obrazu.



Kožní léze jsou morfologicky různorodé, tj. můžeme pozorovat léze tvarově symetrické i asymetrické, barevně homogenní i nehomogenní, s abruptivními i pozvolnými okraji. Pokud jde o barvu lézí, soustředí se model na generování převážně světlehnědých, tmavohnědých a občas červených lézí, které tvoří pouze 3 z 6 kategorií barev lézí podle algoritmu ABCD dle Stolze (dalšími jsou bílá, modrošedá a černá)[77]. Hlavním nedostatkem modelu je absence složitějších morfologických struktur a detailů, jež mají pro diagnostiku zásadní význam. Při hodnocení věrohodnosti obrazu je možné poukázat také na neschopnost modelu syntetizovat prvky přirozené pro okolí léze, jako je například ochlupení.

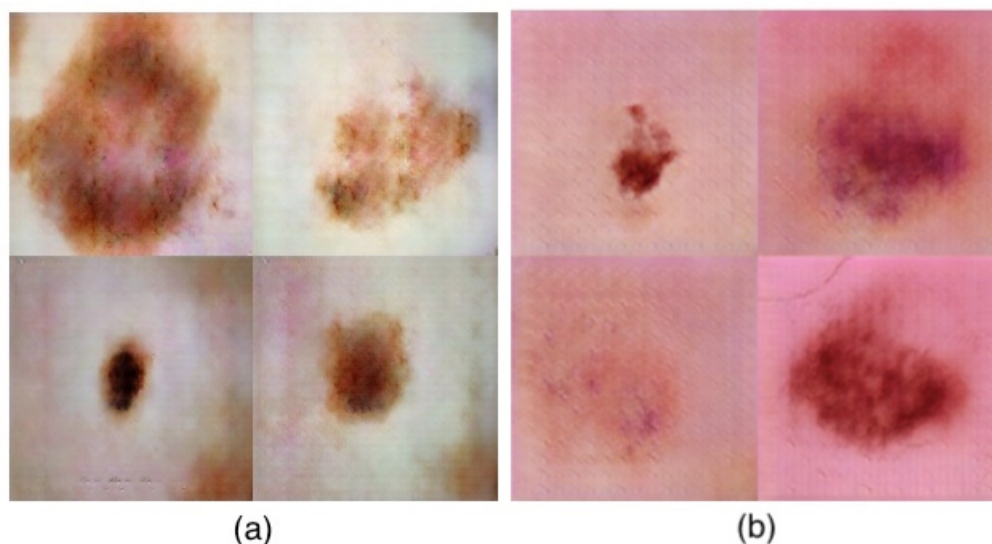
Kvalita vygenerovaných obrazů je různá a pro srovnání s realistickými výstupy na obrázku 7.1 jsou na obrázku 7.2 ukázány méně zdařilé obrazy. Takto výrazně vyčnívající vzorky jsou v poměru s ostatními generovány v poměru asi 1 : 30.



Obr. 7.2: Ukázka nerealistických obrazů kožních lézí generovaných modelem DCGAN. Některé obrazy obsahují léze s nejasnými obrysy (a), bílé skvrny (b) nebo výrazně rušivé barvy (c).

Pro srovnání lze výstupy porovnat s obrazy kožních lézí prezentovanými v rámci publikace Baur et al. [64]. Přestože byl každý model trénován na jiných experimentálních datech, jedná se o data pocházející ze stejného zdroje (archiv „ISIC Challenge Datasets“ [78]) a v obou případech byla pro generování obrazů použita síť typu DCGAN. Z obrázku 7.3 lze usoudit, že kvalita vygenerovaných vzorků je srovnatelná. Použitím konvolučních filtrů s velikostí jádra, jež je mocninou stride (tj. `filter size = (4,4)` a `stride = (2,2)`), se navíc podařilo potlačit šachovnicové architekty typické pro model GAN, na něž bylo v publikaci Baur et al. poukazováno.

Z těchto výsledků lze vyvodit závěr, že model DCGAN je schopen generovat jednoduché medicínské obrazy s omezeným rozlišením a úrovní detailu. Právě detaily struktur zachycených na snímku jsou však při využití dermatoskopických snímků v diagnostice stěžejní a model DCGAN je proto v těchto ohledech nedostačující.



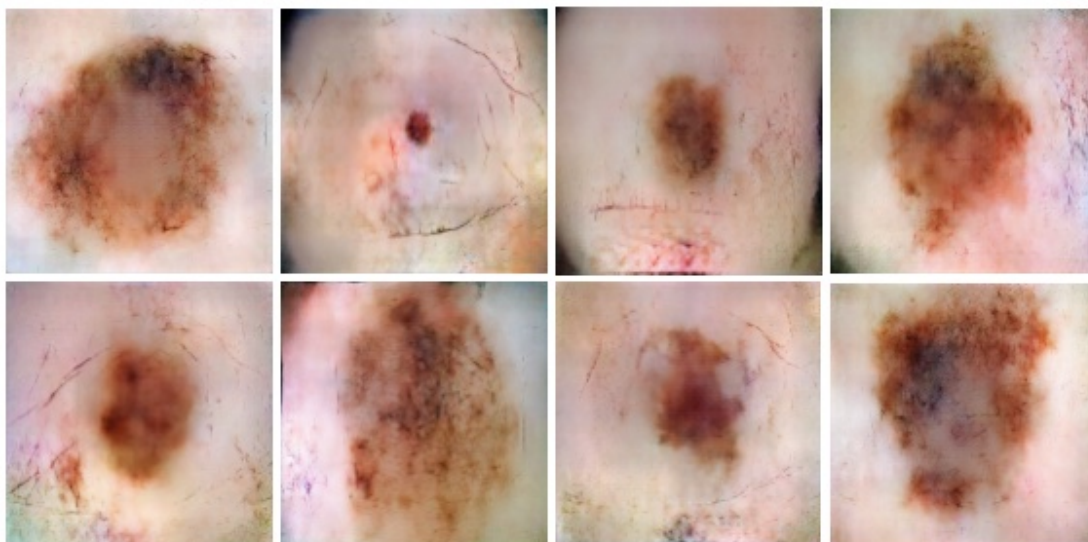
Obr. 7.3: Ukázka syntetických obrazů vygenerovaných modelem sítě typu DCGAN implementovaným v této práci (a) a v rámci publikace Baur et al. (b)) [64].

Dostatečně kvalitní syntetické obrazy kožních lézí lze využít např. pro rozšíření datových sad zobrazujících nedostatečně zastoupené třídy. Pro cílené generování lézí určitých tříd a charakteristik je však třeba generátoru poskytnout další vstupní informace, což vyžaduje použití modelů schopných podmíněné syntézy. Z podmíněných modelů se velmi efektivním ukázal model *pix2pix* [79], jež prokázal schopnost generovat obrazy kožních lézí s vysokým rozlišením ( $1024 \times 512$  pixelů), které kvalitou předčí obrazy generované ostatními modely GAN.

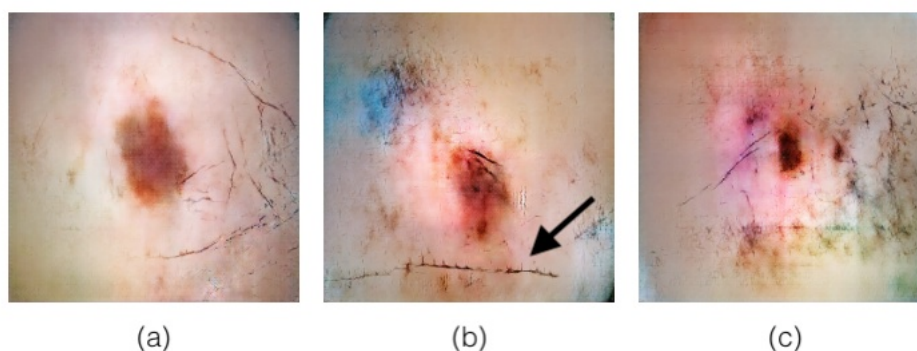
## 7.2 Model „pix2pix“ pro generování obrazů kožních lézí

Vlastní implementace modelu *pix2pix* pro generování obrazů kožních lézí byla popsána v podkapitole 6.2.4. Po úspěšném tréninku byl tento model testován na datové sadě binárních masek kožních lézí. Poskytnutím doplňujících informací modelu byly překonány hlavní nedostatky modelu DCGAN a výsledkem jsou syntetické obrazy kožních lézí výrazně lépe napodobující dermatoskopické snímky. Ukázka těchto výstupů je na obrázku 7.4.

Dalším pokrokem oproti modelu DCGAN je syntéza prvků v okolí kožní léze. Na obrázku 7.5 je příklad výstupů, u nichž se model pokoušel o generování ochlupení a měřítka využívaného při dermatoskopii. U některých obrazů se opět objevují barevné artefakty, jež se nepodařilo zcela potlačit ani v tomto modelu.



Obr. 7.4: Ukázka syntetických obrazů vygenerovaných modelem *pix2pix*. Kožní léze mají realistické zbarvení a detailní strukturu.

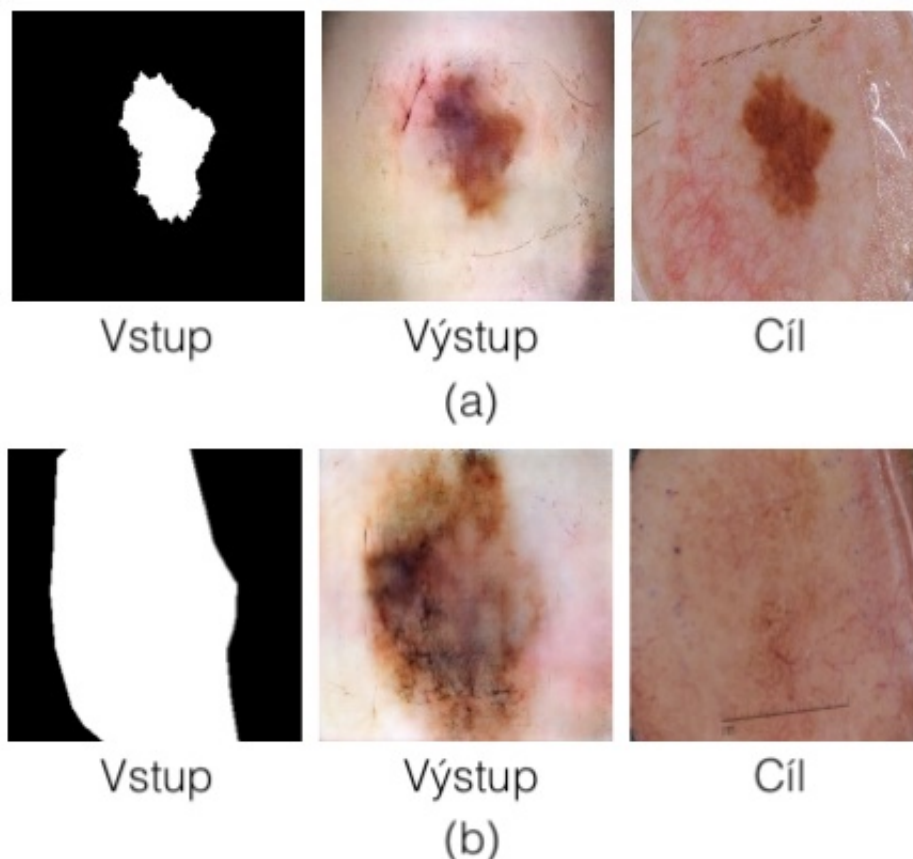


Obr. 7.5: Pokusy generátoru zakomponovat do obrazu ochlupení (a) a měřítko využívané při dermatoskopii (b). Příklad obrazu zatíženého artefakty (c).

Dalším cílem modelu *pix2pix* je na základě binární masky generovat obrazy identické s reálnými snímky. I v tomto ohledu dosahuje model uspokojivých výsledků, avšak pouze při generování určitých typů lézí. Na obrázku 7.6 je porovnán výstup generátoru se skutečným snímkem pro dva případy. Generování kožní léze se světle-hnědým zbarvením proběhlo úspěšně, lézi bez pigmentu však model nebyl schopen realisticky syntetizovat.

Model se obecně zaměřuje na syntézu lézí s benigními znaky a není schopen vytvořit složitější struktury, jež se vyskytují převážně u maligních lézí, jako je např. pigmentová síť, radiální proužky, hnědé globule, černé tečky, bělavý závoj a šedomodré

plošky [77]. Neschopnost modelu tyto struktury syntetizovat může být zapříčiněna malým zastoupením různých maligních tříd lézí v tréninkových datech.



Obr. 7.6: Porovnání výstupu modelu *pix2pix* při generování léze světlehnědé barvy (a) a léze bez pigmentového zbarvení (b).

Lepších výsledků s modelem *pix2pix* by kromě změn architektury a hyperparametrů sítě mohlo být dosaženo také modifikací vstupních informací. Proces podmíněné syntézy lze v případě modelu *pix2pix* regulovat pomocí vstupních binárních masek. Binární masky mohou být vybrány z testovací sady podle požadovaných parametrů nebo mohou být uměle vytvořeny. Pro ukázkou tohoto přístupu byl navržen experiment, jež propojuje nepodmíněný model DCGAN a podmíněný model *pix2pix*. Výsledky tohoto experimentu jsou popsány v podkapitole 7.2.1. Binární maska definující tvar léze však zůstává jediným vstupem modelu.

Pro generování obrazů realistických kožních lézí, jež zároveň nesou charakteristiky určité třídy je třeba generátoru poskytnout další informace. Například v již citované publikaci Bissoto et al. [79] byly generovány obrazy kožních lézí s klinickým významem na základě sémantických anotovaných map (obrázků, v nichž každá hodnota pixelu představuje třídu) a tzv. „instančních map“ (obrázků, v nichž jsou



v pixelech zakódovány informace o třídě a instanci objektu). Takovýto přístup ovšem vyžaduje využití značně modifikovaného modelu *pix2pix* („pix2pix HD“) a náročnou přípravu vstupních dat.

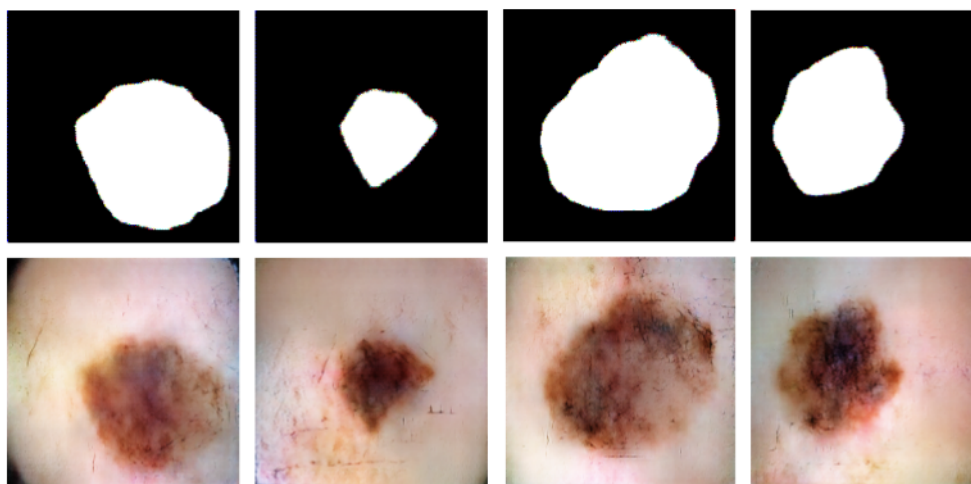
Dalším obdobným příkladem je modifikace modelu *pix2pix* pojmenovaná „DermGAN“ [80], jež generuje vysoce kvalitní snímky kožních lézí na základě sémantické mapy. Vstupní sémantická mapa obsahuje informace o barvě kůže, diagnóze kožního stavu a jeho lokaci na snímku.

I přes některé nedostatky se však *pix2pix* GAN osvědčil při generování medicínských obrazů a v další části práce byl aplikován na komplexnější medicínská data.

### 7.2.1 Experiment se syntetickými maskami

Model DCGAN, implementovaný podle 6.1, byl trénován po dobu 10 učících epoch na datové sadě 2000 binárních masek kožních lézí popsané v 4.1.2. Natrénovaný model byl poté využit ke generování syntetických masek s rozměry  $256 \times 256 \times 3$ . Výstupy modelu DCGAN byly upraveny tak, aby jejich formát přesně odpovídal reálným binárním maskám a byly použity jako vstup modelu *pix2pix*, popsaného v podkapitole 6.2.4.

Ukázka výsledků experimentu je na obrázku 7.7. Modelem DCGAN se již po krátkém tréninku podařilo generovat binární masky vizuálně neodlišitelné od reálných a model *pix2pix* dokázal na základě těchto masek syntetizovat realistické obrazy kožních lézí. Tímto způsobem je možné rozšířit testovací sadu binárních masek pro model *pix2pix*, jenž je na těchto vstupech závislý.



Obr. 7.7: V prvním řádku je ukázka syntetických masek vygenerovaných modelem DCGAN a v druhém řádku k nim odpovídající obrazy kožních lézí vygenerované modelem *pix2pix*.

## 7.3 Model „pix2pix“ pro generování obrazů páteře

Další zkoumanou aplikací modelu *pix2pix* bylo generování obrazů páteře. Pro tento účel byly trénovány a testovány dva samostatné modely, první pro generování obrazů z axiální roviny a druhý z roviny sagitální. Způsob rozdělení dat do trénovacích a testovacích sad je popsán v 4.2 a detaily implementace modelů a tréninkové procesy v 6.2.

Každý z modelů byl trénován po dobu 200 učících epoch a následně byl použit pro generování obrazů na základě masek z testovací datové sady. Výstupy generátoru byly pro účely vyhodnocení zobrazeny ve dvou různých radiologických oknech. První radiologické okno se středem 400 HU a šířkou okna 1000 HU zvýrazňuje v obraze kostní tkáň, zatímco druhé okno se středem 50 HU a šířkou okna 250 HU zvýrazňuje měkkou tkáň. Pokud není uvedeno jinak, na obrázcích je použito radiologické okno pro kost.

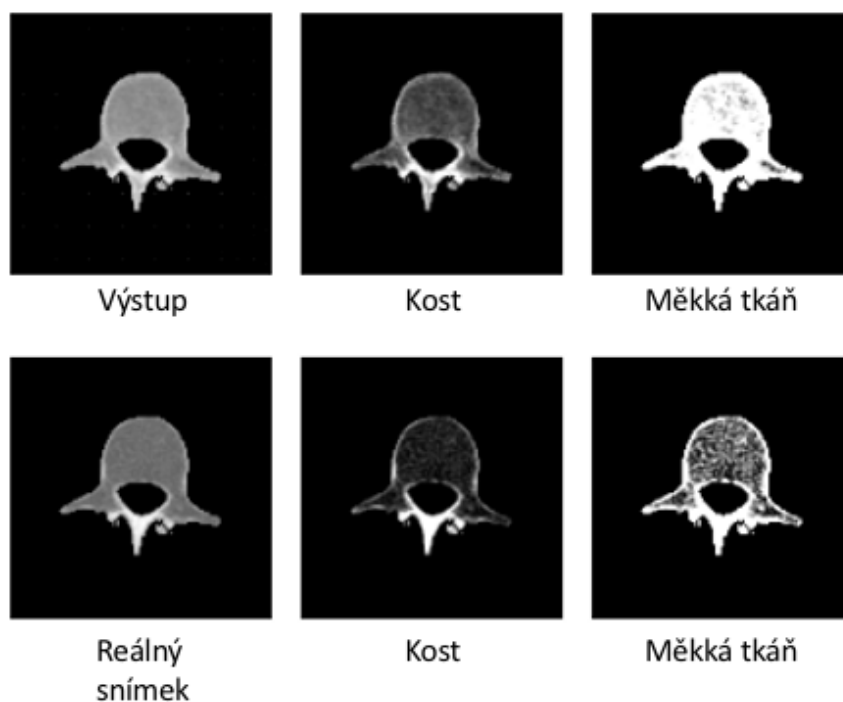
### 7.3.1 Generování axiálních řezů páteře

Prvním kritériem, jímž lze hodnotit modely *pix2pix*, je jejich schopnost aproximovat na základě masky reálný snímek. Na CT snímcích se výrazně projevují patologické stavy páteře, jimiž jsou pacienti postihnuti. Patologické stavy je pro model s takto omezenou datovou sadou obtížné předpovědět. Na obrázku 7.8 je příklad vygenerovaného obrazu bederního obratle ve srovnání s reálným snímkem. Přestože se obrazy zcela neshodují, můžeme na vygenerovaném obraze pozorovat charakteristickou strukturu kosti obratle, blíže popsanou na obrázku 7.9. Generátor správně vyplnil tělo obratle homogenní spongiózní kostní tkání, obratel ohraničil vrstvou kortikální kosti a v oblasti oblouku obratle vygeneroval kost s nejvyšší hustotou. Tyto skutečnosti prokazují, že se generativní síť naučila charakteristické rysy axiálního snímku obratle a je schopna je reprodukovat.

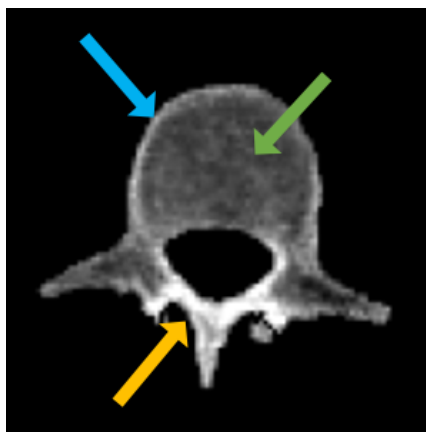
Schopnost modelu generovat na základě vstupu realistický výstup, jež je druhým kritériem hodnocení, byla dále zkoumána za použití různých vstupních masek. Testovací datové sady axiálních snímků od 3 různých pacientů obsahují jak masky fyziologického, tak i patologického tvaru. Předpokladem pro úspěšný model je, aby byl schopen rozpoznat druh masky a na základě této informace generoval obraz obratle s odpovídajícími charakteristikami.

Výsledky testování modelu generujícího axiální řezy potvrdily, že generátor implementovaného modelu tyto předpoklady z velké části opravdu splňuje. Na obrázku 7.10 a 7.11 je ukázka syntetických axiálních snímků, které toto tvrzení podporují.

Obrazy obratlů vygenerované na základě masky s fyziologickým tvarem vykazují znaky zdravých obratlů. Tělo obratle je tvořeno homogenní spongiózní tkání nízké



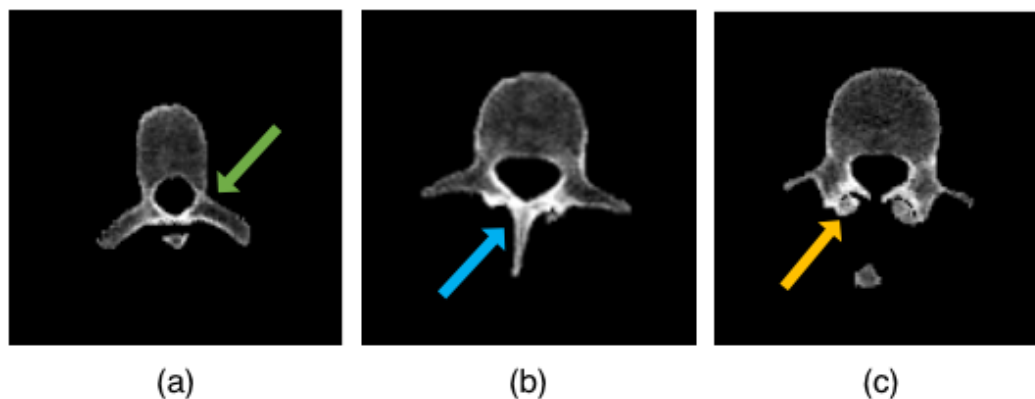
Obr. 7.8: Srovnání vygenerovaného a reálného snímku bederního obratle. V prvním sloupci jsou snímky zobrazeny v plném rozsahu hodnot, v druhém pouze radiologické okno pro kost a ve třetím radiologické okno pro měkkou tkáň.



Obr. 7.9: Vygenerovaný obraz bederního obratle zobrazený v radiologickém okně pro kost. Můžeme pozorovat spongiózní kostní tkáň (zelená šipka), kortikální kostní tkáň (modrá šipka) a výrazný oblouk obratle (žlutá šipka).

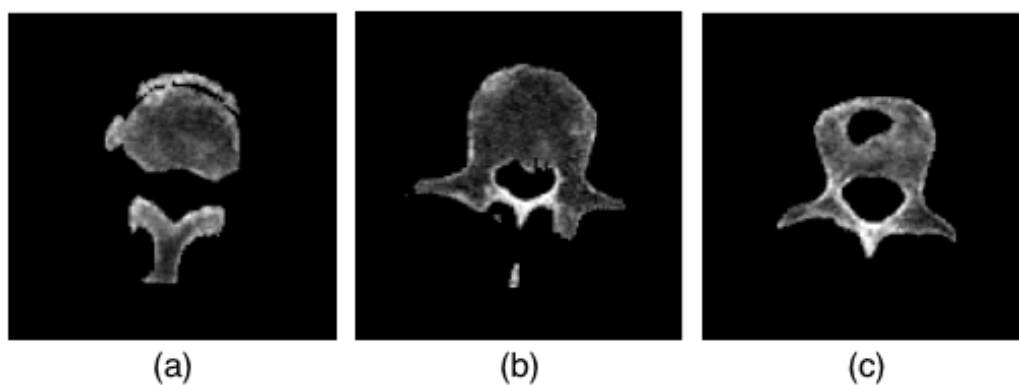
hustoty s nižšími pixelovými hodnotami než zbytek obratle, díky čemuž se na snímku jeví tmavší. Spongiózní část obratle je pokryta tenkou kortikální kostní tkání, jež je na snímku viditelná jako světlá tenká linie lemující obratel. U všech tří obratlů

můžeme pozorovat větší hustotu kosti v oblasti zúžené ploténky, která dorzálním směrem pokračuje v oblouk obratle, taktéž s vysokou hustotou kosti. Světle zobrazená kortikální tkáň se dále nachází v okolí výběžků obratle, které jsou na snímcích vyznačeny u bederních obratlů.



Obr. 7.10: Ukázka obratlů vygenerovaných na základě masky s fyziologickým tvarem. (a) hrudní obratel s vyznačeným zúžením ploténky, (b) bederní obratel s trnovým výběžkem, (c) bederní obratel s vyznačeným kloubním výběžkem.

Za masky s patologickým tvarem jsou považovány masky s nekompaktním tvarem a výraznými erozemi, značícími patologický stav obratle. Model v těchto případech správně generuje nehomogenní spongiózní tkáň, ve níž je porušena trabekulární struktura. Zvýšená hustota kosti se na snímku projevuje světlou barvou a je patrná ve všech ukázkových obrazech.

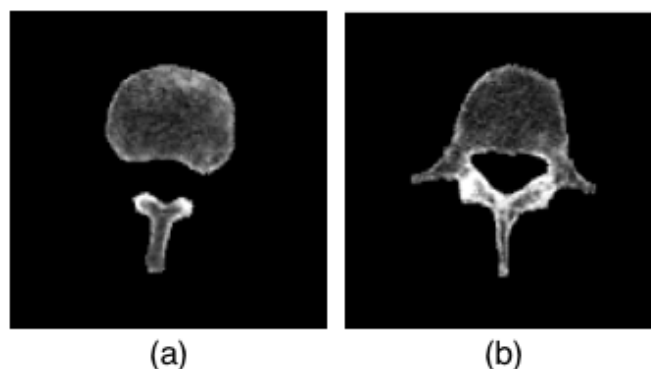


Obr. 7.11: Ukázka obratlů vygenerovaných na základě masky s patologickým tvarem.

V některých případech se modelu nepodařilo správně vyhodnotit vstupní masku. Na obrázku 7.12 lze vidět selhání modelu, kdy generoval bederní obratel s fyziologic-



kou maskou jako patologický a naopak obratel s patologickou maskou jako obratel bez výraznějších patologií.



Obr. 7.12: Ukázka výstupů generátoru s chybnými charakteristikami.

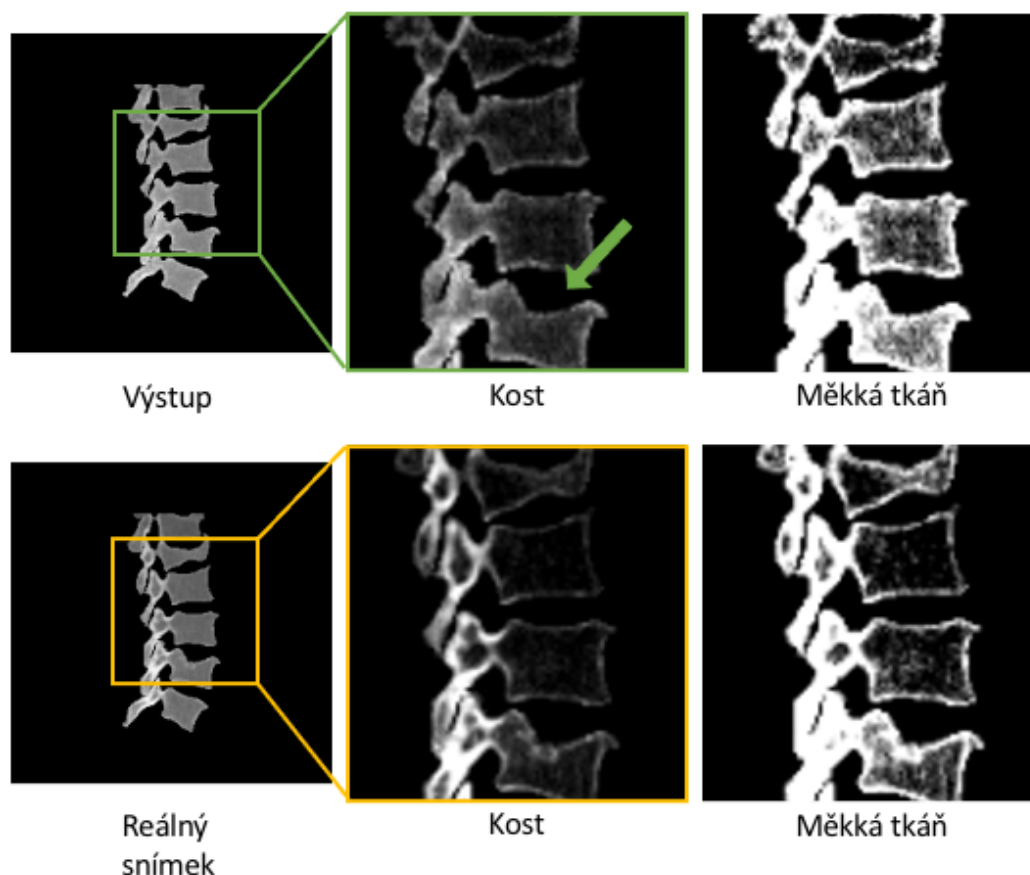
K selhání modelu však během testování došlo pouze v několika případech a lze tedy usoudit, že generátor modelu *pix2pix* je schopen generovat realistické obrazy axiálních CT řezů páteře.

### 7.3.2 Generování sagitálních řezů páteře

Jako další byl hodnocen model *pix2pix* pro generování sagitálních řezů páteře. Testování tohoto modelu probíhalo na datové sadě zahrnující sagitální CT snímky bederní páteře 3 pacientů. Pro porovnání syntetických a reálných snímků je na obrázku 7.13 přiblížen jeden z výstupů generátoru a odpovídající reálný snímek.

Obdobně jako u axiálních řezů i u generovaných sagitálních řezů můžeme pozorovat základní charakteristiky snímků obratlů. Každý z obratlů je lemován tenkou vrstvou světlé kortikální kosti. Vyšší pixelové hodnoty má v obraze zúžení mezi tělem a obloukem obratle a oblast trnových výběžků, což značí vyšší hustotu kosti v těchto místech. Vygenerované obrazy však vykazují i viditelné nedostatky. Model nedokázal syntetizovat trabekulární strukturu spongiózní tkáně obratlů, kterou lze vidět u reálného snímku a jeden ze zborcených obratlů byl modelem interpretován jako fyziologický.

Přestože architektura a tréninkový proces obou modelů pro generování obrazů páteře jsou obdobné, pozorujeme výrazně jinou kvalitu výstupů. Zásadním rozdílem mezi modely je velikost a rozmanitost tréninkových dat. Tréninková datová sada sagitálních snímků zahrnuje pouze 389 vzorků, což se pro naučení modelu ukázalo jako nedostačující. Model však i přes takto omezené množství vstupních dat demonstroval schopnost reprodukovat základní charakteristiky reálných snímků, což může být považováno za alespoň dílčí úspěch.



Obr. 7.13: Srovnání vygenerovaného a reálného snímku bederního páteře. V prvním sloupci jsou snímky zobrazeny v plném rozsahu hodnot, v druhém pouze radiologické okno pro kost a ve třetím radiologické okno pro měkkou tkáň. Zelenou šipkou je označen nesprávně generovaný zborcený obratel.

## 7.4 Diskuze na téma generování medicínských obrazů modely GAN

Obsahem této kapitoly je shrnutí poznatků o generování medicínských obrazů pomocí modelů GAN.

### 7.4.1 Nepodmíněná syntéza medicínských obrazů

První zkoumanou aplikací GAN byla nepodmíněná syntéza medicínských obrazů na základě vstupního vektoru šumu. Model DCGAN byl schopen syntetizovat obrazy kožních lézí s rozlišením  $256 \times 256$ , a přestože by dalšími experimenty se sítí bylo jistě možné ještě zvýšit kvalitu výstupního obrazu, možnosti tohoto modelu jsou značně omezené. Model DCGAN je vhodný pro seznámení se s modely GAN, prvky

využívanými v jejich architektuře a faktory ovlivňujícími stabilitu tréninku. Pro nepodmíněnou syntézu se však nabízejí vhodnější modifikace.

První z nich je model Wasserstein GAN, jež je možné implementovat několika úpravami modelu DCGAN. Zásadní modifikací je využití alternativní kritériální funkce, konkrétně Wassersteinovy vzdálenosti. Pro modely s touto kritériální funkcí se snáze určují hyperparametry, modely vykazují větší stabilitu během tréninku, je eliminován kolaps režimu generátoru a pomocí kritériální funkce je zároveň možné i objektivní hodnocení kvality vygenerovaných snímků [81]. Hodnoty vah diskriminátoru modelu WGAN jsou omezeny na fixní interval (často  $\langle -0.01, 0.01 \rangle$ ), který musí být pečlivě zvolen, aby trénink modelu konvergoval. Alternativou tohoto modelu je WGAN s penalizací gradientu (z angl. „Wasserstein GAN Gradient Penalty“, WGAN-GP), jež místo prostého prahování hodnot vah model penalizuje v případě, že se normy gradientů příliš odchýlí od hodnoty 1 [82]. Modelem WGAN a jeho modifikací WGAN-GP lze generovat vysoce kvalitní medicínské obrazy s klinickým významem [83, 84] a zároveň mají oproti dále uvedeným progresivním modelům výrazně jednodušší implementaci. Wassersteinova vzdálenost a Wassersteinova vzdálenost s penalizací gradientu mohou být též využity jako kritériální funkce dalších modifikací modelu GAN s cílem stabilizace tréninku [85].

Další modely vhodné pro přímou syntézu snímků ze šumu jsou již zmíněné progresivní modely LAPGAN a PGAN. Model LAPGAN je omezen na generování realistických obrazů pouze s nízkým rozlišením, jež bude obdobné jako u DCGAN pro většinu medicínských aplikací nedostačující. Naopak model PGAN prokázal schopnost generovat medicínské snímky s rozlišením až  $1024 \times 1024$  [40, 64]. Syntetické obrazy vygenerované modelem PGAN mohou usnadnit trénování klasifikačních hlubokých neuronových sítí na malých datových sadách medicínských snímků, a tím zlepšit výsledky klasifikace [86].

#### 7.4.2 Podmíněná syntéza medicínských obrazů

Podmíněná syntéza medicínských obrazů byla zkoumána pomocí modelu *pix2pix*, jež generuje výstup na základě vstupního obrazu. „Image-to-image“ transformace je hlavním předmětem podmíněné syntézy medicínských obrazů a většina publikací na toto téma používá pro takovou syntézu modifikace právě tohoto modelu. S dostatečně velkou tréninkovou datovou sadou je základní model *pix2pix* schopen generovat vysoce realistické medicínské obrazy. V některých publikacích je diskutován vznik šachovnicového artefaktu ve výstupním obraze jako důsledek použití vrstev transponované konvoluce v architektuře sítě [80]. Kvůli nízkému rozlišení výstupních obrazů v této práci nebyl tento artefakt nijak výrazný a proto mu nebyla věnována větší pozornost. Při generování obrazů s vyšším rozlišením je možné vrstvu

transponované konvoluce nahradit nadvzorkovací vrstvou `Upsampling2D` následovanou konvoluční vrstvou `Conv2D`, jak je vysvětleno v kapitole popisující implementaci modelu DCGAN.

Implementovaný model by mohl dosahovat dobrých výsledků v oblasti intermodální syntézy obrazu, kdy je vstupem sítě např. MRI snímek a na jeho základě je vygenerován odpovídající snímek CT. V tomto případě není síti třeba poskytovat další vstupní informace. Naopak pro ostatní typy „Image-to-image“ transformace vyžaduje architektura a učící proces modelu *pix2pix* výraznější modifikace navržené s ohledem na konkrétní aplikaci.

# Závěr

Bakalářská práce se zaměřuje na problematiku generování syntetických obrazů s využitím strojového učení se zaměřením na generativní soutěživé sítě. Hlavním cílem teoretické části bylo seznámení se s problematikou generativně soutěživých sítí a vypracování rešerše na publikované metody a využití těchto sítí v medicíně. Cílem praktické části bylo dle rešerše navrhnout využití generativních soutěživých sítí pro medicínské účely a implementovat řešení ve vhodném programovém prostředí. Výsledky generování medicínských obrazů za použití implementovaných modelů bylo dále nutno zhodnotit a diskutovat.

V teoretickém úvodu byly obecně popsány umělé neuronové sítě se zaměřením na konvoluční neuronové sítě a především na generativní soutěživé sítě. Konkrétně druhá kapitola se věnuje popisu generativně soutěživých sítí, jejich mnoha modifikacím a využití těchto sítí v medicíně. Jsou vybrány tři publikace na téma generování medicínských obrazů s využitím těchto sítí a tyto aplikace jsou ve třetí kapitole blíže popsány. Čtvrtá kapitola se věnuje popisu a předzpracování použitých medicínských datových sad, mezi něž patří dvě sady dermatoskopických snímků a CT snímky páteře 12 anonymních pacientů. Další kapitoly se zabývají návrhem modelu hluboce konvoluční generativní soutěživé sítě a modelu „pix2pix“ pro generování medicínských obrazů a implementací těchto modelů v programovém prostředí Python s využitím knihoven Tensorflow a Keras.

V sedmé kapitole jsou výsledky generování medicínských obrazů modely generativních soutěživých sítí vyhodnoceny na základě subjektivního vizuálního dojmu. Model hluboce konvoluční generativně soutěživé sítě byl schopen generovat obrazy kožních lézí srovnatelné s výsledky vybrané vědecké publikace, ovšem stále s omezeným rozlišením a úrovní detailu obrazu. Řada nedostatků hluboce konvolučního modelu byla překonána modelem *pix2pix*, jenž na základě binární masky dokázal syntetizovat obrazy kožních lézí s realistickým zbarvením, detailní strukturou a v omezené kvalitě také ochlupením v okolí léze. Obrazy kožních lézí syntetizované v této práci však postrádají složitější struktury stěžejní pro diagnostický význam těchto obrazů, na základě čehož je doporučeno rozšíření modelu. Model *pix2pix* dále prokázal schopnost generovat na základě binární masky realistické obrazy axiálních CT řezů páteře a s mírnými nedostatky taktéž obrazy sagitálních CT řezů páteře. V případě komplexních medicínských dat, jako jsou CT snímky, je nutné poskytnout modelu dostatečné množství tréninkových dat.

Využití generativních soutěživých sítí pro generování medicínských obrazů má velký potenciál. Modifikace těchto modelů dosahují skvělých výsledků při syntéze vysoce kvalitních medicínských obrazů, díky čemuž mohou být využity pro mnoho aplikací v oblasti medicíny.

# Literatura

- [1] YEGNANARAYANA, Bayya. *Artificial Neural Networks*. India: Prentice-Hall of India Pvt.Ltd, 2004. ISBN 9788120312531.
- [2] The structure of a real neuron. In: *ResearchGate* [online]. [cit. 2020-12-20]. Dostupné z: [https://www.researchgate.net/profile/Zafeirios\\_Fountas/publication/266485234/figure/fig1/AS:651919710228482@1532441179600/The-structure-of-a-real-neuron.png](https://www.researchgate.net/profile/Zafeirios_Fountas/publication/266485234/figure/fig1/AS:651919710228482@1532441179600/The-structure-of-a-real-neuron.png)
- [3] JAIN, A.K., JIANCHANG MAO a K.M. MOHIUDDIN. Artificial neural networks: a tutorial. *Computer* [online]. 1996, **29**(3), 31-44 [cit. 2020-12-10]. ISSN 00189162. Dostupné z: doi:10.1109/2.485891
- [4] GRAUPE, Daniel. *Principles of Artificial Neural Networks*. 3rd edition. World Scientific Publishing Company, 2013. ISBN 978-9814522731.
- [5] Matematický model a aktivní dynamika neuronu. In: *Matematická biologie učebnice* [online]. [cit. 2020-12-20]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=analýza-a-hodnocení-biologických-dat-umělá-inteligence-neuronové-site-jednotlivý-neuron-jednotlivý-neuron-matematický-model-a-aktivní-dynamika-neuronu>
- [6] NOVÁK, Mirko. *Umělé neuronové sítě: teorie a aplikace* [online]. Praha: C.H. Beck, 1998 [cit. 2020-12-19]. Učebnice informatiky. ISBN isbn80-7179-132-6. Dostupné z: <https://ndk.cz/view/uuid:0fe12ff0-306f-11e9-b81e-005056827e52?page=uuid:9f6443f0-5d61-11e9-ace0-005056827e51>
- [7] Artificial neural network architecture. In: *ResearchGate* [online]. [cit. 2020-12-20]. Dostupné z: [https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o\\_fig1\\_321259051](https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051)
- [8] MALÝ, Marek. *Vícevrstvé dopředné neuronové sítě: úvod do teorie a aplikací*. Ústí nad Labem: Univerzita J.E. Purkyně, Přírodovědecká fakulta, 2007. ISBN isbn978-80-7044-915-8.
- [9] GOODFELLOW, Ian, Yoshua BENGIO a Aaron COURVILLE. *Deep learning* [online]. Cambridge, MA: MIT press, [2016]. [cit. 2021-5-25]. Adaptive computation and machine learning series. ISBN 9780262035613. Dostupné z: <https://www.deeplearningbook.org>

- [10] ROJAS, Raúl. The Backpropagation Algorithm. ROJAS, Raúl. *Neural Networks* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, 1996, s. 149-182 [cit. 2021-5-25]. ISBN 978-3-540-60505-8. Dostupné z: doi:10.1007/978-3-642-61068-4\_7
- [11] KINGMA, Diederik P. a Jimmy BA. *Adam: A Method for Stochastic Optimization* [online]. 2014 [cit. 2021-01-02]. Dostupné z: <https://arxiv.org/abs/1412.6980>
- [12] Architecture of CNN. *Research Gate* [online]. [cit. 2021-01-01]. Dostupné z: [https://www.researchgate.net/figure/Architecture-of-the-CNN-used\\_fig2\\_303303279](https://www.researchgate.net/figure/Architecture-of-the-CNN-used_fig2_303303279)
- [13] *Princip výpočtu dvourozměrné diskrétní konvoluce.* [online]. 19 July 2006 [cit. 2021-5-19]. Dostupné z: [https://commons.wikimedia.org/wiki/File:Konvoluce\\_2rozm\\_diskretni.jpg](https://commons.wikimedia.org/wiki/File:Konvoluce_2rozm_diskretni.jpg)
- [14] YAMASHITA, Rikiya, Mizuho NISHIO, Richard Kinh Gian DO a Kaori TOGASHI. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging* [online]. 2018, **9**(4), 611-629 [cit. 2020-12-19]. ISSN 1869-4101. Dostupné z: doi:10.1007/s13244-018-0639-9
- [15] *How to use LeakyReLU as an Activation Function in Keras?* [online]. 2020 [cit. 2021-01-03]. Dostupné z: <https://androidkt.com/how-to-use-leakyrelu-as-an-activation-function-in-keras/>
- [16] KIM, Phil. Convolutional Neural Network. KIM, Phil. *MATLAB Deep Learning* [online]. Berkeley, CA: Apress, 2017, 2017-06-16, s. 121-147 [cit. 2020-12-19]. ISBN 978-1-4842-2844-9. Dostupné z: doi:10.1007/978-1-4842-2845-6\_6
- [17] ALBAWI, Saad, Tareq Abed MOHAMMED a Saad AL-ZAWI. Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)* [online]. IEEE, 2017, 2017, s. 1-6 [cit. 2020-12-19]. ISBN 978-1-5386-1949-0. Dostupné z: doi:10.1109/ICEngTechnol.2017.8308186
- [18] *A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way* [online]. 2018 [cit. 2021-01-03]. Dostupné z: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- [19] DUMOULIN, Vincent a Francesco VISIN. *A guide to convolution arithmetic for deep learning* [online]. 2016 [cit. 2021-5-24]. Dostupné z: <https://arxiv.org/abs/1603.07285>

- [20] IOFFE, Sergey a Christian SZEGEDY. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* [online]. [cit. 2021-5-24]. Dostupné z: <https://arxiv.org/abs/1502.03167>
- [21] SRIVASTAVA, Nitish, Geoffrey HINTON, Alex KRIZHEVSKY, Ilya SUTSKEVER a Ruslan SALAKHUTDINOV. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* [online]. **2014**(15), 19291958 [cit. 2021-5-24]. ISSN 1929-1958. Dostupné z: <https://www.semanticscholar.org/paper/Dropout%3A-a-simple-way-to-prevent-neural-networks-Srivastava-Hinton/34f25a8704614163c4095b3ee2fc969b60de4698>
- [22] The Dropout Layer. *Baeldung* [online]. [cit. 2021-5-24]. Dostupné z: <https://www.baeldung.com/cs/ml-relu-dropout-layers>
- [23] GOODFELLOW, Ian J., Jean POUGET-ABADIE, Mehdi MIRZA, Bing XU, David WARDE-FARLEY, Sherjil OZAIR, Aaron COURVILLE a Yoshua BEN-GIO. *Generative Adversarial Nets* [online]. 2014 [cit. 2021-01-02]. Dostupné z: <https://arxiv.org/pdf/1406.2661.pdf>
- [24] *Vanilla GAN with Numpy* [online]. 2019 [cit. 2021-01-03]. Dostupné z: <https://christinakouridi.blog/2019/07/09/vanilla-gan-numpy/>
- [25] *Generative Adversarial Networks: Build Your First Models* [online]. 2020 [cit. 2021-01-03]. Dostupné z: <https://realpython.com/generative-adversarial-networks/>
- [26] *A Gentle Introduction to Generative Adversarial Networks (GANs)* [online]. 2019 [cit. 2021-01-03]. Dostupné z: <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>
- [27] PIETERS, Mathijs a Marco WIERING. *Comparing Generative Adversarial Network Techniques for Image Creation and Modification* [online]. 2018 [cit. 2021-01-02]. Dostupné z: <https://arxiv.org/abs/1803.09093>
- [28] KAZEMINIA, Salome, Christoph BAUR, Arjan KUIJPER, Bram VAN GINNEKEN, Nassir NAVAB, Shadi ALBARQOUNI a Anirban MUKHOPADHYAY. GANs for medical image analysis. *Artificial Intelligence in Medicine* [online]. 2020, **109** [cit. 2020-12-10]. ISSN 09333657. Dostupné z: [doi:10.1016/j.artmed.2020.101938](https://doi.org/10.1016/j.artmed.2020.101938)
- [29] YI, Xin, Ekta WALIA a Paul BABYN. Generative adversarial network in medical imaging: A review. *Medical Image Analysis* [online]. 2019, **58** [cit. 2020-12-10]. ISSN 13618415. Dostupné z: [doi:10.1016/j.media.2019.101552](https://doi.org/10.1016/j.media.2019.101552)



- [30] ZHANG, Qianqian, Haifeng WANG, Hongya LU, Daehan WON a Sang Won YOON. Medical Image Synthesis with Generative Adversarial Networks for Tissue Recognition. In: *2018 IEEE International Conference on Healthcare Informatics (ICHI)* [online]. New York, NY, USA: IEEE, 2018, 2018, s. 199-207 [cit. 2020-12-10]. ISBN 978-1-5386-5377-7. ISSN 2575-2634. Dostupné z: doi:10.1109/ICHI.2018.
- [31] RADFORD, Alec, Luke METZ a Soumith CHINTALA. *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks* [online]. 2015 [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1511.06434>
- [32] MIRZA, Mehdi a Simon OSINDERO. *Conditional Generative Adversarial Nets* [online]. 2014 [cit. 2020-12-30]. Dostupné z: <http://arxiv.org/abs/1411.1784>
- [33] ISOLA, Phillip, Jun-Yan ZHU, Tinghui ZHOU a Alexei A. EFROS. Image-to-Image Translation with Conditional Adversarial Networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. Honolulu, HI, USA: IEEE, 2017, 2017, s. 5967-5976 [cit. 2020-12-10]. ISBN 978-1-5386-0457-1. ISSN 1063-6919. Dostupné z: doi:10.1109/CVPR.2017.632
- [34] CHEN, Xi, Yan DUAN, Rein HOUTHOOFT, John SCHULMAN, Ilya SUTSKEVER a Pieter ABBEEL. *InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets* [online]. 2016 [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1606.03657>
- [35] DENTON, Emily, Soumith CHINTALA, Arthur SZLAM a Rob FERGUS. Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks. *Advances in neural information processing systems* [online]. 2015, , 1486—1494 [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1506.05751>
- [36] ZHANG, Han, Tao XU, Hongsheng LI, Shaoting ZHANG, Xiaogang WANG, Xiaolei HUANG a Dimitris METAXAS. StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks. In: *2017 IEEE International Conference on Computer Vision (ICCV)* [online]. Venice, Italy: IEEE, 2017, 2017, s. 5908-5916 [cit. 2020-12-30]. ISBN 978-1-5386-1032-9. ISSN 2380-7504. Dostupné z: doi:10.1109/ICCV.2017.629
- [37] KARRAS, Tero, Timo AILA, Samuli LAINE a Jaakko LEHTINEN. *Progressive Growing of GANs for Improved Quality, Stability, and Variation* [online]. [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1710.10196>
- [38] CHUQUICUSMA, Maria J. M., Sarfaraz HUSSEIN, Jeremy BURT a Ulas BAGCI. How to fool radiologists with generative adversarial networks? A visual

- turing test for lung cancer diagnosis. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* [online]. Washington, DC, USA: IEEE, 2018, 2018, s. 240-244 [cit. 2020-12-30]. ISBN 978-1-5386-3636-7. ISSN 1945-8452. Dostupné z: doi:10.1109/ISBI.2018.8363564
- [39] FRID-ADAR, Maayan, Eyal KLANG, Michal AMITAI, Jacob GOLDBERGER a Hayit GREENSPAN. Synthetic data augmentation using GAN for improved liver lesion classification. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* [online]. IEEE, 2018, 2018, s. 289-293 [cit. 2020-12-30]. ISBN 978-1-5386-3636-7. Dostupné z: doi:10.1109/ISBI.2018.8363576
- [40] BEERS, Andrew, James BROWN, Ken CHANG, J. Peter CAMPBELL, Susan OSTMO, Michal F. a Jayashree KALPATHY-CRAMER. *High-resolution medical image synthesis using progressively grown generative adversarial networks* [online]. 2018 [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1805.03144>
- [41] CALIMERI, Francesco, Aldo MARZULLO, Claudio STAMILE a Giorgio TERRACINA. Biomedical Data Augmentation Using Generative Adversarial Neural Networks. LINTAS, Alessandra, Stefano ROVETTA, Paul F.M.J. VERSCHURE a Alessandro E.P. VILLA, ed. *Artificial Neural Networks and Machine Learning — ICANN 2017* [online]. Cham: Springer International Publishing, 2017, 2017-10-25, s. 626-634 [cit. 2020-12-30]. Lecture Notes in Computer Science. ISBN 978-3-319-68611-0. Dostupné z: doi:10.1007/978-3-319-68612-7\_71
- [42] JIN, Dakai, Ziyue XU, Youbao TANG, Adam P. HARRISON a Daniel J. MOLLURA. CT-Realistic Lung Nodule Simulation from 3D Conditional Generative Adversarial Networks for Robust Lung Segmentation. FRANGI, Alejandro F., Julia A. SCHNABEL, Christos DAVATZIKOS, Carlos ALBEROLA-L-PEZ a Gabor FICHTINGER, ed. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2018* [online]. Cham: Springer International Publishing, 2018, 2018-09-26, s. 732-740 [cit. 2020-12-30]. Lecture Notes in Computer Science. ISBN 978-3-030-00933-5. Dostupné z: doi:10.1007/978-3-030-00934-2\_81
- [43] COSTA, Pedro, Adrian GALDRAN, Maria Ines MEYER, Meindert NIEMEIJER, Michael ABRAMOFF, Ana Maria MENDONCA a Aurelio CAMPILHO. End-to-End Adversarial Retinal Image Synthesis. *IEEE Transactions on Medical Imaging* [online]. 2018, **37**(3), 781-791 [cit. 2020-12-30]. ISSN 0278-0062. Dostupné z: doi:10.1109/TMI.2017.2759102

- [44] MAHAPATRA, Dwarikanath, Behzad BOZORGTABAR, Jean-Philippe THIRAN a Mauricio REYES. Efficient Active Learning for Image Classification and Segmentation Using a Sample Selection and Conditional Generative Adversarial Network. FRANGI, Alejandro F., Julia A. SCHNABEL, Christos DAVATZIKOS, Carlos ALBEROLA-L-PEZ a Gabor FICHTINGER, ed. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2018* [online]. Cham: Springer International Publishing, 2018, 2018-09-26, s. 580-588 [cit. 2020-12-30]. Lecture Notes in Computer Science. ISBN 978-3-030-00933-5. Dostupné z: doi:10.1007/978-3-030-00934-2\_65
- [45] JIN, Cheng-Bin, Hakil KIM, Mingjie LIU, et al. Deep CT to MR Synthesis Using Paired and Unpaired Data. *Sensors* [online]. 2019, **19**(10) [cit. 2020-12-30]. ISSN 1424-8220. Dostupné z: doi:10.3390/s19102361
- [46] HUO, Yuankai, Zhoubing XU, Hyeonsoo MOON, et al. SynSeg-Net: Synthetic Segmentation Without Target Modality Ground Truth. *IEEE Transactions on Medical Imaging* [online]. 2019, **38**(4), 1016-1025 [cit. 2020-12-30]. ISSN 0278-0062. Dostupné z: doi:10.1109/TMI.2018.2876633
- [47] ZHAO, Yu, Shu LIAO, Yimo GUO, et al. Towards MR-Only Radiotherapy Treatment Planning: Synthetic CT Generation Using Multi-view Deep Convolutional Neural Networks. FRANGI, Alejandro F., Julia A. SCHNABEL, Christos DAVATZIKOS, Carlos ALBEROLA-L-PEZ a Gabor FICHTINGER, ed. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2018* [online]. Cham: Springer International Publishing, 2018, 2018-09-26, s. 286-294 [cit. 2021-5-27]. Lecture Notes in Computer Science. ISBN 978-3-030-00927-4. Dostupné z: doi:10.1007/978-3-030-00928-1\_33
- [48] BEN-COHEN, Avi, Eyal KLANG, Stephen P. RASKIN, Shelly SOFFER, Simona BEN-HAIM, Eli KONEN, Michal Marianne AMITAI a Hayit GREENSPAN. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. *Engineering Applications of Artificial Intelligence* [online]. 2019, **78**, 186-194 [cit. 2020-12-30]. ISSN 09521976. Dostupné z: doi:10.1016/j.engappai.2018.11.013
- [49] PAN, Yongsheng, Mingxia LIU, Chunfeng LIAN, Tao ZHOU, Yong XIA a Dinggang SHEN. Synthesizing Missing PET from MRI with Cycle-consistent Generative Adversarial Networks for Alzheimer-s Disease Diagnosis. FRANGI, Alejandro F., Julia A. SCHNABEL, Christos DAVATZIKOS, Carlos ALBEROLA-L-PEZ a Gabor FICHTINGER, ed. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2018* [online]. Cham:

- Springer International Publishing, 2018, 2018-09-13, s. 455-463 [cit. 2020-12-30]. Lecture Notes in Computer Science. ISBN 978-3-030-00930-4. Dostupné z: doi:10.1007/978-3-030-00931-1\_52
- [50] WOLTERINK, Jelmer M., Tim LEINER, Max A. VIERGEVER a Ivana ISGUM. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. *IEEE Transactions on Medical Imaging* [online]. 2017, **36**(12), 2536-2545 [cit. 2020-12-10]. ISSN 0278-0062. Dostupné z: doi:10.1109/TMI.2017.2708987
- [51] CHEN, Yuhua, Feng SHI, Anthony G. CHRISTODOULOU, Yibin XIE, Zhengwei ZHOU a Debiao LI. Efficient and Accurate MRI Super-Resolution Using a Generative Adversarial Network and 3D Multi-level Densely Connected Network. FRANGI, Alejandro F., Julia A. SCHNABEL, Christos DAVATZIKOS, Carlos ALBEROLA-L-PEZ a Gabor FICHTINGER, ed. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2018* [online]. Cham: Springer International Publishing, 2018, 2018-09-26, s. 91-99 [cit. 2020-12-10]. Lecture Notes in Computer Science. ISBN 978-3-030-00927-4. Dostupné z: doi:10.1007/978-3-030-00928-1\_11
- [52] WANG, Yan, Biting YU, Lei WANG, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *NeuroImage* [online]. 2018, **174**, 550-562 [cit. 2020-12-10]. ISSN 10538119. Dostupné z: doi:10.1016/j.neuroimage.2018.03.045
- [53] HU, Bo, Ye TANG, Eric I-Chao CHANG, Yubo FAN, Maode LAI a Yan XU. Unsupervised Learning for Cell-Level Visual Representation in Histopathology Images With Generative Adversarial Networks. *IEEE Journal of Biomedical and Health Informatics* [online]. 2019, **23**(3), 1316-1328 [cit. 2020-12-10]. ISSN 2168-2194. Dostupné z: doi:10.1109/JBHI.2018.2852639
- [54] CHEN, Dali, Yingying AO a Shixin LIU. Semi-Supervised Learning Method of U-Net Deep Learning Network for Blood Vessel Segmentation in Retinal Images. *Symmetry* [online]. 2020, **12**(7) [cit. 2020-12-10]. ISSN 2073-8994. Dostupné z: doi:10.3390/sym12071067
- [55] RASHID, Haroon, M. Asjid TANVEER a Hassan AQEEL KHAN. Skin Lesion Classification Using GAN based Data Augmentation. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* [online]. Berlin, Germany: IEEE, 2019, 2019, s. 916-919 [cit. 2020-12-10]. ISBN 978-1-5386-1311-5. ISSN 1558-4615. Dostupné z: doi:10.1109/EMBC.2019.8857905

- [56] MADANI, Ali, Mehdi MORADI, Alexandros KARARGYRIS a Tanveer SYEDA-MAHMOOD. Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* [online]. Washington, DC, USA: IEEE, 2018, 2018, s. 1038-1042 [cit. 2020-12-10]. ISBN 978-1-5386-3636-7. ISSN 1945-8452. Dostupné z: doi:10.1109/ISBI.2018.8363749
- [57] YANG, Dong, Daguang XU, S. Kevin ZHOU, Bogdan GEORGESCU, Mingqing CHEN, Sasa GRBIC, Dimitris METAXAS a Dorin COMANICIU. *Automatic Liver Segmentation Using an Adversarial Image-to-Image Network* [online]. 2017 [cit. 2020-12-30]. Dostupné z: <https://arxiv.org/abs/1707.08037>
- [58] SCHLEGL, Thomas, Philipp SEEBÖCK, Sebastian M. WALDSTEIN, Ursula SCHMIDT-ERFURTH a Georg LANGS. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. NIETHAMMER, Marc, Martin STYNER, Stephen AYLWARD, Hongtu ZHU, Ipek OGUZ, Pew-Thian YAP a Dinggang SHEN, ed. *Information Processing in Medical Imaging* [online]. Cham: Springer International Publishing, 2017, 2017-05-23, s. 146-157 [cit. 2020-12-10]. Lecture Notes in Computer Science. ISBN 978-3-319-59049-3. Dostupné z: doi:10.1007/978-3-319-59050-9\_12
- [59] STYNER, Martin A., Elsa D. ANGELINI, Varghese ALEX, Mohammed SAFWAN K. P., Sai Saketh CHENNAMSETTY a Ganapathy KRISHNA-MURTHI. Generative adversarial networks for brain lesion detection. In: *Proceedings of the SPIE* [online]. Volume 10133. Orlando, Florida, United States: SPIE, 2017, 2017-2-24, 101330G- [cit. 2020-12-10]. ISBN 978-1-7281-0831-5. Dostupné z: doi:10.1117/12.2254487
- [60] HU, Yipeng, Eli GIBSON, Tom VERCAUTEREN, Hashim U. AHMED, Mark EMBERTON, Caroline M. MOORE, J. Alison NOBLE a Dean C. BARRATT. Intraoperative Organ Motion Models with an Ensemble of Conditional Generative Adversarial Networks. DESCOTEAUX, Maxime, Lena MAIER-HEIN, Alfred FRANZ, Pierre JANNIN, D. Louis COLLINS a Simon DUCHESNE, ed. *Medical Image Computing and Computer-Assisted Intervention MICCAI 2017* [online]. Cham: Springer International Publishing, 2017, 2017-09-04, s. 368-376 [cit. 2020-12-30]. Lecture Notes in Computer Science. ISBN 978-3-319-66184-1. Dostupné z: doi:10.1007/978-3-319-66185-8\_42
- [61] ROSS, Tobias, David ZIMMERER, Anant VEMURI, et al. Exploiting the potential of unlabeled endoscopic video data with self-supervised learning.

- International Journal of Computer Assisted Radiology and Surgery* [online]. 2018, **13**(6), 925-933 [cit. 2020-12-30]. ISSN 1861-6410. Dostupné z: doi:10.1007/s11548-018-1772-0
- [62] MAHMOOD, Rafid, Aaron BABIER, Andrea MCNIVEN, Adam DIAMANT a Timothy C. Y. CHAN. Automated Treatment Planning in Radiation Therapy using Generative Adversarial Networks. *Proceedings of the 3rd Machine Learning for Healthcare Conference* [online]. 2018, (85), 484-499 [cit. 2020-12-30].
- [63] GOODFELLOW, Ian. *NIPS 2016 Tutorial: Generative Adversarial Networks* [online]. 2016 [cit. 2020-12-31]. Dostupné z: <http://arxiv.org/abs/1701.00160>
- [64] BAUR, Christoph, Shadi ALBARQOUNI a Nassir NAVAB. Generating Highly Realistic Images of Skin Lesions with GANs. STOYANOV, Danail, Zeike TAYLOR, Duygu SARIKAYA, et al., ed. *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis* [online]. Cham: Springer International Publishing, 2018, 2018-10-02, s. 260-267 [cit. 2020-12-10]. Lecture Notes in Computer Science. ISBN 978-3-030-01200-7. Dostupné z: doi:10.1007/978-3-030-01201-4\_28
- [65] ARGENZIANO, Giuseppe a H Peter SOYER. Dermoscopy of pigmented skin lesions — a valuable tool for early diagnosis of melanoma. *The Lancet Oncology* [online]. 2001, **2**(7), 443-449 [cit. 2021-5-15]. ISSN 14702045. Dostupné z: doi:10.1016/S1470-2045(00)00422-8
- [66] *ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection* [online]. 2018 [cit. 2020-12-31]. Dostupné z: <https://challenge2018.isic-archive.com>
- [67] GALBUSERA, Fabio, Frank NIEMEYER, Maike SEYFRIED, Tito BASSANI, Gloria CASAROLI, Annette KIENLE a Hans-Joachim WILKE. Exploring the Potential of Generative Adversarial Networks for Synthesizing Radiological Images of the Spine to be Used in In Silico Trials. *Frontiers in Bioengineering and Biotechnology* [online]. 2018, **6** [cit. 2020-12-20]. ISSN 2296-4185. Dostupné z: doi:10.3389/fbioe.2018.00053
- [68] *In silico Clinical Trials: How Computer Simulation will Transform the Biomedical Industry* [online]. In: . Brussels, Belgium, 2015 [cit. 2020-12-20]. Dostupné z: [https://avicenna-alliance.com/files/user\\_upload/PDF/Avicenna\\_Roadmap.pdf](https://avicenna-alliance.com/files/user_upload/PDF/Avicenna_Roadmap.pdf)
- [69] NIE, Dong, Roger TRULLO, Jun LIAN, Caroline PETITJEAN, Su RUAN, Qian WANG a Dinggang SHEN. Medical Image Synthesis with Context-Aware

- Generative Adversarial Networks. DESCOTEAUX, Maxime, Lena MAIERHEIN, Alfred FRANZ, Pierre JANNIN, D. Louis COLLINS a Simon DUCHESNE, ed. *Medical Image Computing and Computer Assisted Intervention MICCAI 2017* [online]. Cham: Springer International Publishing, 2017, 2017-09-04, s. 417-425 [cit. 2020-12-31]. Lecture Notes in Computer Science. ISBN 978-3-319-66178-0. Dostupné z: doi:10.1007/978-3-319-66179-7\_48
- [70] ROTEMBERG, Veronica, Nicholas KURTANSKY, Brigit BETZ-STABLEIN, et al. A Patient-Centric Dataset of Images and Metadata for Identifying Melanomas Using Clinical Context. *The ISIC 2020 Challenge Dataset* [online]. 2020 [cit. 2020-12-30]. Dostupné z: <https://challenge2020.isic-archive.com>
- [71] CODELLA, N, D GUTMAN, ME CELEBI, et al. Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC). *ISIC Challenge Datasets* [online]. [cit. 2021-5-15]. Dostupné z: <https://challenge.isic-archive.com/data>
- [72] SEN, Hrithick. *Understanding GANs - Deriving the Adversarial loss from scratch* [online]. [cit. 2021-5-18]. Dostupné z: <https://medium.com/analytics-vidhya/understanding-gans-deriving-the-adversarial-loss-from-scratch-ccd8b683d7e2>
- [73] RONNEBERGER, Olaf, Philipp FISCHER a Thomas BROX. U-Net: Convolutional Networks for Biomedical Image Segmentation. NAVAB, Nassir, Joachim HORNEGGER, William M. WELLS a Alejandro F. FRANGI, ed. *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015* [online]. Cham: Springer International Publishing, 2015, 2015-11-18, s. 234-241 [cit. 2021-5-17]. Lecture Notes in Computer Science. ISBN 978-3-319-24573-7. Dostupné z: doi:10.1007/978-3-319-24574-4\_28
- [74] *Welcome to Python* [online]. [cit. 2021-5-18]. Dostupné z: <https://www.python.org>
- [75] Train Generative Adversarial Network (GAN). *MathWorks* [online]. [cit. 2020-12-30]. Dostupné z: <https://www.mathworks.com/help/deeplearning/ug/train-generative-adversarial-network.html>
- [76] SHMELKOV, Konstantin, Cordelia SCHMID a Karteek ALAHARI. How good is my GAN? *Proceedings of the European Conference on Computer Vision (ECCV)* [online]. **2018**, pp. 213-229 [cit. 2021-5-20]. Dostupné z: <https://arxiv.org/abs/1807.09499>

- [77] DRLÍK, Lubomír, Veronika PAŤAVOVÁ a Lumír POCK. Dermatoskopie. *Dermatol. praxi* [online]. 2014, (8(3)), 102—107 [cit. 2021-5-20]. Dostupné z: <https://www.dermatologiepropraxi.cz/pdfs/der/2014/03/05.pdf>
- [78] *ISIC Challenge Datasets* [online]. [cit. 2021-01-03]. Dostupné z: <https://challenge.isic-archive.com/data>
- [79] BISSOTO, Alceu, Fábio PEREZ, Eduardo VALLE a Sandra AVILA. *Skin Lesion Synthesis with Generative Adversarial Networks* [online]. 2019 [cit. 2021-01-03]. Dostupné z: <https://arxiv.org/abs/1902.03253>
- [80] GHORBANI, Amirata, Vivek NATARAJAN, COZ a Yuan LIU. DermGAN: Synthetic Generation of Clinical Skin Images with Pathology. *Proceedings of the Machine Learning for Health NeurIPS Workshop* [online]. 2020, (116), 155-170 [cit. 2021-5-25]. Dostupné z: <http://proceedings.mlr.press/v116/ghorbani20a.html>
- [81] ARJOVSKY, Martin, Soumith CHINTALA a Léon BOTTOU. *Wasserstein GAN* [online]. 2017 [cit. 2021-01-03]. Dostupné z: <https://arxiv.org/abs/1701.07875>
- [82] GULRAJANI, Ishaan, Faruk AHMED, Martin ARJOVSKY, Vincent DUMOULIN a Aaron C. COURVILLE. Improved Training of Wasserstein GANs. *Advances in Neural Information Processing Systems 30* [online]. 2017 [cit. 2021-5-25]. Dostupné z: <https://papers.nips.cc/paper/2017/hash/892c3b1c6dcd52936e27cbd0ff683d6-Abstract.html>
- [83] HAN, Changhee, Hideaki HAYASHI, Leonardo RUNDO, et al. GAN-based synthetic brain MR image generation. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* [online]. IEEE, 2018, 2018, s. 734-738 [cit. 2021-5-25]. ISBN 978-1-5386-3636-7. Dostupné z: doi:10.1109/ISBI.2018.8363678
- [84] LEI, Ke, Morteza MARDANI, John M. PAULY a Shreyas S. VASANAWALA. Wasserstein GANs for MR Imaging: From Paired to Unpaired Training. *IEEE Transactions on Medical Imaging* [online]. 2021, **40**(1), 105-115 [cit. 2021-5-25]. ISSN 0278-0062. Dostupné z: doi:10.1109/TMI.2020.3022968
- [85] KWON, Gihyun, Chihye HAN a Dae-shik KIM. *Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks* [online]. [cit. 2021-5-25]. Dostupné z: <https://arxiv.org/abs/1908.02498>



- [86] ZHANG, Huijuan, Zongrun HUANG a Zhongwei LV. Medical Image Synthetic Data Augmentation Using GAN. In: *Proceedings of the 4th International Conference on Computer Science and Application Engineering* [online]. New York, NY, USA: ACM, 2020, 2020-10-20, s. 1-6 [cit. 2021-5-25]. ISBN 9781450377720. Dostupné z: doi:10.1145/3424978.3425118
- [87] NIE, Dong, Roger TRULLO, Jun LIAN, Li WANG, Caroline PETITJEAN, Su RUAN, Qian WANG a Dinggang SHEN. Medical Image Synthesis with Deep Convolutional Adversarial Networks. *IEEE Transactions on Biomedical Engineering* [online]. 2018, **65**(12), 2720-2730 [cit. 2021-5-25]. ISSN 0018-9294. Dostupné z: doi:10.1109/TBME.2018.2814538

## Seznam zkratek

<b>CNN</b>	Convolutional Neural Network
<b>ReLU</b>	Rectifier Linear Unit
<b>lReLU</b>	Leaky Rectifier Linear Unit
<b>GAN</b>	Generative Adversarial Network
<b>DCGAN</b>	Deep Convolutional Generative Adversarial Network
<b>cGAN</b>	Conditional Generative Adversarial Network
<b>LAPGAN</b>	Laplacian Generative Adversarial Network
<b>StackGAN</b>	Stacked Generative Adversarial Network
<b>PGAN</b>	Progressive Growing Generative Adversarial Network
<b>WGAN</b>	Wasserstein Generative Adversarial Network
<b>WGAN-GP</b>	Wasserstein Generative Adversarial Network Gradient Penalty

# A Obsah elektronických příloh

Součástí práce jsou elektronické přílohy, jež obsahují následující složky a soubory.

```
/ ..... kořenový adresář přiložených souborů
├── DCGAN.....složka souborů náležících k modelu DCGAN
│   ├── DCGAN_preprocess.py.....skript využitý k předzpracování dat
│   ├── testing_demo.....složka se soubory potřebnými k testování modelu
│   │   ├── DCGAN_model_645.h5.....natrénovaný generátor modelu DCGAN
│   │   ├── DCGAN_testing_demo.py.....skript určený k testování modelu DCGAN
│   │   └── demo_output.....složka pro uložení výstupů testování
│   └── training.....ukázkové skripty pro trénování modelu DCGAN
│       ├── DCGAN_data_loader.py.....pomocné funkce - načítání dat
│       ├── DCGAN_models_256.py.....pomocné funkce - definice modelu DCGAN
│       └── DCGAN.py.....hlavní skript pro trénování modelu DCGAN
├── other_models.....složka s dalšími modely neprezentovanými v práci
│   ├── DCGAN_models_64.py.....definice modelu DCGAN s rozlišením 64 × 64
│   ├── DCGAN_models_128.py.....definice modelu DCGAN s rozlišením 128 × 128
│   ├── PGAN.py.....skript pro trénování modelu PGAN
│   └── WGAN.py.....skript pro trénování modelu WGAN
├── pix2pix_melanoma.....složka souborů náležících k modelu pix2pix - kožní léze
│   ├── pix2pix_melanoma_preprocess.py.....skript využitý k předzpracování dat
│   ├── testing_demo.....složka se soubory potřebnými k testování modelu
│   │   ├── DCGAN_melanoma_model_010.h5 .. natrénovaný generátor modelu DCGAN
│   │   ├── demo_output.....složka pro uložení výstupů testování
│   │   ├── pix2pix_melanoma_model_200.h5 . natrénovaný generátor modelu pix2pix
│   │   └── pix2pix_testing_demo.py.....skript určený k testování modelu pix2pix
│   └── training.....ukázkové skripty pro trénování modelu pix2pix
│       ├── pix2pix_melanoma_data_loader.py.....pomocné funkce - načítání dat
│       ├── pix2pix_melanoma_models.py..pomocné funkce - definice modelu pix2pix
│       └── pix2pix_melanoma.py.....hlavní skript pro trénování modelu pix2pix
├── pix2pix_spine.....složka souborů náležících k modelu pix2pix - páteř
│   ├── nifti_preprocess.py.....skript využitý k předzpracování dat
│   ├── pix2pix_axial.py..složka souborů náležících k modelu pix2pix - axiální řezy
│   │   ├── testing_demo.....složka se soubory potřebnými k testování modelu
│   │   │   ├── demo_output.....složka pro uložení výstupů testování
│   │   │   ├── model_axial_200.h5.....natrénovaný generátor modelu pix2pix
│   │   │   ├── pix2pix_axial_input_demo.npz..soubor se vstupy pro model pix2pix
│   │   │   └── pix2pix_testing_demo.py...skript určený k testování modelu pix2pix
│   │   └── training.....ukázkové skripty pro trénování modelu pix2pix
│   │       ├── pix2pix_axial_data_loader.py.....pomocné funkce - načítání dat
│   │       ├── pix2pix_axial_models.py..pomocné funkce - definice modelu pix2pix
│   │       └── pix2pix_axial.py.....hlavní skript pro trénování modelu pix2pix
│   └── pix2pix_sagital.py...složka souborů náležících k modelu pix2pix - sagitální řezy
│       ├── testing_demo.....složka se soubory potřebnými k testování modelu
│       └── demo_output.....složka pro uložení výstupů testování
```

- model\_sagital\_200.h5 ..... natrénovaný generátor modelu pix2pix
  - pix2pix\_sagital\_input\_demo.npz soubor se vstupy pro model pix2pix
  - pix2pix\_testing\_demo.py...skript určený k testování modelu pix2pix
- training ..... ukázkové skripty pro trénování modelu pix2pix
  - pix2pix\_sagital\_data\_loader.py.....pomocné funkce - načítání dat
  - pix2pix\_sagital\_models.py pomocné funkce - definice modelu pix2pix
  - pix2pix\_sagital.py ..... hlavní skript pro trénování modelu pix2pix
- README.txt ..... pokyny k používání elektronických příloh
- requirements.txt ..... požadavky na knihovny a jejich verze